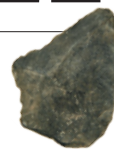


# THIS WEEK

## EDITORIALS

**NEUROSCIENCE** End of the beginning for US brain-mapping project **p.6**

**WORLD VIEW** Water subsidies hamper the fight against drought **p.7**



**STONE AGE** Ancient flaky rock-tool technique had multiple births **p.9**

## Retraction challenges

*Cleaning up the literature can be difficult.*

A key responsibility of any journal is to correct erroneous information that it has published, and as quickly as possible.

Easily said! It is straightforward enough for authors to correct a paper. But if it becomes clear after publication that the conclusions are fundamentally flawed, a retraction is appropriate — and things can then get a lot more challenging.

Why, other than through enforcement after misconduct, would anyone retract a paper in a high-profile journal? Regrettably, given the reputational damage that a retraction might yield, it may take a strong code of honour, and a strong consensus among sometimes many co-authors, to go public, rather than just let the paper join the many others that turn out to be flawed and fade away.

That is why the literature of retractions in high-impact journals might be skewed towards misconduct that has been proved through investigations. But all praise to the authors who decide to behave honourably. Where authors make it clear that nothing more than an honest error was involved, their retraction should bring them credit.

Where misconduct — a deliberate attempt to deceive — has been involved, things tend to get complicated. Universities fear misconduct for the immense trouble that it can cause them in investigations, for the legal tussles that can then ensue if the proceedings are contested, and for the potential damage to their reputations. But when such investigations prove misconduct, they often lead to retractions of one or many papers. Even then, if the conclusions are contested, journals might find themselves threatened with a lawsuit for the proposed retraction itself, let alone a retraction whose statement includes any reference to misconduct.

For years, with occasional exceptions, *Nature's* annual number of research-paper retractions tended to average around one or two. But over the past two years, we have seen a considerable rise — six in 2013, and seven, so far, in 2014. We have reviewed these and previous retractions and would like to make some observations on the basis of their content and on the experiences of publishing them.

A high proportion of *Nature's* retractions in recent years have come about through honest error, where authors have either discovered mistakes themselves after publication, or have had the errors brought to their attention and taken action.

Another observation is that negotiating some retractions can involve unavoidable delays of years because of some combination of the complexity of the science, disputes between co-authors, the need to await outcomes of lengthy investigations, and disputes over these proceedings. Journal editors have neither the authority nor the means to police authors or their institutions, and can be dependent on proceedings whose details are confidential to institutions. They also need to be sensitive to the interests of blameless co-authors.

Even when an institution and a journal both want a retraction, their interests in doing so may collide. An institution might be bound by confidentiality agreements and therefore unable to release the results of its scientific investigations, leaving editors in the dark as to the

circumstances behind erroneous work. An institution may also wish the wording of the retraction to bolster its case against a wrong-doer, whereas a journal's interest is to avoid lengthy disputes, push the paper into oblivion, and avoid further wasted effort by researchers. Whether for that reason or, occasionally, for legal reasons, we have concluded that we cannot usually use retraction statements as a means of highlighting wrong-doing.

Why the sudden pulse of *Nature* retractions in 2013 and 2014? (The last year to reach such heights was 2003, when we retracted seven fraudulent papers by the physicist Jan Hendrik Schön.) We can only speculate. The publication dates of the papers retracted in the past two years range from 1994 to 2014. Data are nowadays more openly

***"The duty to retract a demonstrably false paper remains paramount."***

available and online scrutiny is increasingly vigorous. Some of the rise may parallel the growth in formal corrections associated with increased problems of irreproducibility, which in turn can arise from sloppiness in some overly pressurized laboratories.

That should add to the concern of those worried about wasted funds for research.

But the concerned should also pay attention to what must be increasing costs in legal fees, because those under investigation increasingly turn to lawyers to defend themselves and their reputations, and their employers and journals are more frequently having to respond accordingly. But whatever the obstacles, the duty to retract a demonstrably false paper remains paramount. ■

## Warming up

*Prospects for international agreement on combating climate change look brighter.*

There is much for the world to be pessimistic about these days. The double crises of the Ebola outbreak in West Africa and Islamic extremism in the Middle East, for example, pose real dangers. So it says much for the one-day United Nations summit on climate change, held in New York City last week, that not only did it receive widespread media coverage, but also the enduring message sent by the meeting was one of optimism.

There have been enough 'turning points' in the politics of the effort to curb global warming to send anyone dizzy. That is the narrative the story demands: incremental progress is boring; grand gestures are preferred. Every meeting and announcement is the most important, at least since the previous one.

The politics and the science of climate change have long since parted

company. The science demands political action to aggressively curb greenhouse-gas emissions. The politics, as the saying goes, is a bit more complicated than that. But it is politics, not science, that offers the opportunity for intervention. (The science, of course, can help to guide policy, as is explained in a Comment on page 30 on the absurdity of the 2°C target for global temperature rise.)

If last week's meeting marked a political turning point (and these things are best judged from a distance), then it came with the first signs that the world's largest economies (and worst polluters) are at long last forging an alliance. Even so, the message that seeped from many speeches and presentations was sobering: the combustion of fossil fuels that powers mobility and production in the globalized economy and that keeps our homes warm will probably lead to greater climate change than civilization can easily handle. There is no easy way out of that situation. But although time is running out, the world is not yet doomed.

One lesson, at least, does seem to have been learnt. The top-down approach to emissions reductions — binding caps and legally mandated targets for cuts — is a logical response to the climate problem, but an unworkable one. Global warming is a real and omnipresent risk, but it proceeds slowly and is essentially unobservable to the general public. Unlike escalating epidemics or savage acts of terrorism, a shifting climate has not forced societies and policy-makers to make it a priority. Despite headlines about extreme weather, that is unlikely to change.

Given this political reality, if a new global agreement to tackle climate change is to be agreed by the end of next year — as the UN schedule dictates — then it cannot follow the top-down model of the Kyoto Protocol. But the legal architecture and modes of operation of a contrasting 'bottom-up' agreement remain to be defined. Parties to the United Nations Framework Convention on Climate Change must yet resolve such thorny issues as compliance, verification of reported emissions, and rules of emissions trading. Such technicalities, which often prove to be pitfalls, do matter. But if China, the United States and

the European Union — the world's largest emitters — pull together as they have promised, then a meaningful international climate agreement is well within reach.

Will such an agreement limit warming to 2°C? Almost certainly not. Will it respond appropriately to the scientific evidence of the scale of the likely threat? Definitely not. Is it the best the world can do? Probably.

Regardless of its specifics and legal force, however, a climate agreement will not 'save the world', but nor would a failure of the Paris climate summit in December 2015 automatically mean Armageddon.

The binary rhetoric that campaigners tend to apply in environmental matters does not do justice to the complexity of the task at hand. It would be too easy to blame this or that government for not doing enough when man-made climate change is really the result of collective economic activities, past and present, that cannot be broken like a habit. Key to coming to terms with the unprecedented dilemma we face is effective international cooperation across all aspects of economic and social life, with the ultimate goal of closing the door on the fossil-fuel age.

That goal seems far off, given the continued lure of oil and gas and the huge amount of 'locked-in' emissions from the army of new coal-powered plants in China and elsewhere. And the world population keeps growing: by mid-century, when global emissions will already need to have declined substantially to avoid excessive warming, billions of 'consumers' in Africa and Asia will remain trapped in the fossil-fuel age regardless of the low-carbon technologies that might then be available — unless they are helped out of poverty. Rich countries, meanwhile, must improve their public transport systems, encourage energy-saving construction and invest in grids and energy-storage technology that can accommodate the ebb and flow of electricity from renewable sources. Without these and countless other steps, any climate agreement will ultimately fall short. ■

***"The ultimate goal of closing the door on the fossil-fuel age seems far off."***

## BRAIN gain

*A mixture of focus and innovation is the way forward for big neuroscience.*

As *Nature* went to press, the US National Institutes of Health (NIH) was preparing to announce which scientists it has chosen to help it decipher the brain. To borrow a phrase from Winston Churchill, the announcement could mark the end of the beginning of an effort described by the White House as the greatest since the Human Genome Project. Now all that remains is to unlock the mysteries of the most complex object in the known Universe.

US President Barack Obama announced the BRAIN Initiative (Brain Research through Advancing Innovative Neurotechnologies) 18 months ago. Responsibility for the US\$100-million-a-year project was shared between three agencies: the NIH, the National Science Foundation (NSF) and the Defense Advanced Research Projects Agency (DARPA).

Barely had the initiative fired a synapse before critics attacked its nebulous goal of 'mapping the brain'. Congress had no plans to grant new money for it, and neuroscientists worried that funds would be redirected from other research to support a poorly conceived government mandate. BRAIN's creators enjoyed comparing it with the Human Genome Project, but others drew comparisons with the European Union's Human Brain Project (HBP): a controversial €1-billion (US\$1.3-billion) investment supporting a single researcher's vision of building a computational model of the human brain.

The NIH last year put together a working group to draw up a complex 146-page plan outlining priorities and milestones for BRAIN until 2025.

It did a good job. Although the resulting \$4.5-billion wish list for the project is a tall order, researchers overall are satisfied with an outline for mapping and monitoring the brain that leaves room for innovation.

Fans of both top-down and bottom-up science also got their way. DARPA, with typical military precision, announced that it wanted therapeutic devices for brain disorders that affect soldiers and veterans. It awarded a handful of multimillion-dollar grants to test brain-stimulation systems for purposes such as restoring memory and treating traumatic brain injury. The therapeutic goals are regimented, but the recipients must relish the chance to learn about brain function.

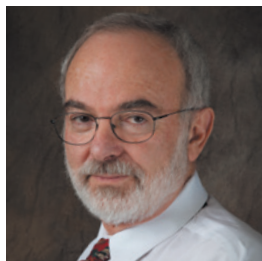
The NSF, which does not normally fund medical and applied research, has taken the opposite approach. In March, it sent out a letter inviting researchers to submit any and all brain-circuit-related ideas as two-page documents. That culminated in a set of 36 small projects, developing everything from tools to image neuron activity to predictive models of brain function. More than any other agency, the NSF shows that big science need not swamp investigator-driven research.

With other brain projects springing up — Israel's investment in brain technologies and Japan's effort to map connections in a marmoset brain, to name just two — the world looked set to form a global collective mind. Then, in July, the HBP derailed less than a year after its launch (see *Nature* 511, 133–134; 2014). Scientists mutinied against director Henry Markram, asking the European Commission to intervene in what they saw as poor management and a focus on simulation rather than neuroscience. As the project struggles with its future, faith in big neuroscience has been shaken and joint HBP–BRAIN plans have been postponed.

The US BRAIN Initiative has the chance to get the concept back on its feet. Success will probably be down to a careful balance between focused order and innovative chaos — much like the organ itself. ■

➔ **NATURE.COM**  
To comment online,  
click on Editorials at:  
[go.nature.com/xhunq](http://go.nature.com/xhunq)





## Water politics must adapt to a warming world

*As rainfall patterns shift, technological and legislative changes are needed to address water shortages, says Moshe Alamaro.*

**T**he world has just experienced the hottest August since records began, and 2014 is shaping up to be one of the warmest years. What will happen to our rainfall as the globe continues to heat up? In theory, a warmer atmosphere should lead to increased ocean evaporation that, in turn, would bring increased precipitation. In practice, many countries have experienced severe drought in recent years — a problem that Brazil and the southwestern United States are currently facing.

Where is the rain? It could be that climate change alters global atmospheric circulation, which leads to increased variability in precipitation. If rainfall patterns are changing, then policies and strategies to conserve water on the ground must shift as well. Long-term planning must change. Attitudes to water must mature. This is not the present situation. Instead, in response to drought we have a series of seemingly futile measures that are best described as a political placebo.

China and several US states, for instance, try to encourage rain by seeding clouds with chemicals launched from aircraft or large guns, despite widespread scepticism about the effectiveness of such measures. And for the past two months in Wichita Falls, Texas, 5 tonnes of a palm oil and limestone powder have been dumped every two days into a 57-square-kilometre reservoir that has shrunk to one-fifth of its capacity because of the drought. The palm oil creates a thin film on the water surface and is claimed to reduce evaporation by 10–30%. But even if it works, a 10% reduction of evaporation is less than the natural year-to-year variability of evaporation due to climate fluctuation. California aims to reduce residential water consumption by 20% through implementing fines on wastage. But residential water use is less than 15% of the total demand, with the rest used mainly for agriculture. Thus a 20% reduction in residential demand will amount to less than 3% of total demand — a mere drop in the bucket.

These measures are the placebo. Governments are compelled to conduct visible (and sometimes noisy, as in the case of cloud-shooting in China) measures to convince their citizens that something is being done — just as police are instructed to use sirens when they drive, to assure people that crime is being fought.

There are better solutions to water shortages. But they will require technological, policy and legislative changes that governments seem unwilling or unable to make. The reality is that there is simply not enough fresh water for everybody to use as much as they want, whenever they want. It makes little sense, for example, to continue to grant ownership — and unlimited use — of underground water to whoever owns the land where the water

is pumped to the surface. Still, at least selling and buying underground water is subject, more or less, to market forces. Surface water from rivers, lakes and reservoirs presents a much more complicated problem.

Water policy is based on voodoo economics — the cost, value and price of water are notoriously difficult to pin down. To an economist, the 'cost' of a litre of surface water in California or Texas is close to zero. Why? Because the reservoirs, dams and aqueducts were built between 50 and 100 years ago, and the investment has long since depreciated.

The 'value' of water is more flexible. In times of drought, governments might cut supply and compensate farmers for the loss of a single growing season. Yet to an almond or olive farmer, the value of one year's water is higher because to grow such a tree takes several seasons. Farmers are therefore willing to pay much more for water during drought.

The 'price' of water, however, is unlike that of most other commodities. Water is not traded between countries — its prices are highly political and subsidized according to the influence of various interest groups. The Santa Clara Valley Water District in San Jose, California, for example, charges US\$40 per acre-foot (about 1.2 million litres) for water for agricultural purposes, and more than \$600 per acre-foot for other industrial applications.

The United States needs a more rational water sector. As climate change continues, it makes little sense, for example, to use heavily subsidized water supplies to grow rice in California or Texas when the crop could be imported from water-rich countries in southeast Asia. (The excessive water, incidentally, is needed not to grow the rice but to suppress weeds.)

To address water shortages, Texas plans to develop new reservoirs. The depreciated cost of construction per acre-foot is more than \$600. This is about the same cost as of desalination. Water-industry reform would open the door to alternative technologies — desalination included — that cannot compete in currently distorted markets. For the United States and others to encourage innovation and ensure access to fresh water, the old system of subsidies must be reformed.

In the second half of the last century, new regulations drastically changed the telecommunications and electricity industries in the United States and elsewhere. This success could be transferred to the water sector worldwide, starting now with federal-guided reforms in drought-stricken California and Texas. ■

**Moshe Alamaro** is a research affiliate in the Department of Earth, Atmospheric, and Planetary Sciences at the Massachusetts Institute of Technology in Cambridge, and chief technology officer of More Aqua, which works on evaporation suppression.  
e-mail: [alamaro@mit.edu](mailto:alamaro@mit.edu)

**THERE IS SIMPLY  
NOT ENOUGH  
FRESH WATER  
FOR EVERYBODY  
TO USE  
AS MUCH AS  
THEY WANT.**

➔ **NATURE.COM**  
Discuss this article  
online at:  
[go.nature.com/t47gvr](http://go.nature.com/t47gvr)

# RESEARCH HIGHLIGHTS

Selections from the  
scientific literature

## NANOTECHNOLOGY

### Microphone made from a molecule

A single molecule can act as a nanometre-sized microphone.

Michel Orrit and his colleagues at the University of Leiden in the Netherlands placed molecules of dibenzoterrylene within a crystal a few degrees above absolute zero and attached a tuning fork to the crystal. Hitting the fork caused vibrations that stretched and compressed the crystal, which in turn shifted the frequency at which the molecules emitted light. The light-frequency readout allowed the team to detect the vibrations from an individual molecule.

The nano-microphone could be used as an ultra-sensitive detector for very slight vibrations, such as from tiny oscillators that measure the properties of quantum systems, the authors say.

*Phys. Rev. Lett.* 113, 135505 (2014)

## NEUROSCIENCE

### Paralysed rats stimulated to walk

Paralysed rats can be made to walk using a device that electrically stimulates the spine and adjusts the pulses according to ongoing movement.

Grégoire Courtine and his



colleagues from the Swiss Federal Institute of Technology in Lausanne implanted electrodes into the spinal cords of rats below the site of the animals' paralyzing injury. The team developed algorithms that tuned the electrical signals in realtime, based on continuous feedback on the leg's position and movement. This allowed the rats to walk with a more natural gait, compared with systems currently in development that use fixed stimulation parameters. The animals walked at least 1,000 steps on a treadmill and could climb steps (pictured).

The authors plan to test their technique in patients with spinal-cord injury. *Sci. Transl. Med.* 6, 255ra133 (2014)

## CANCER

### Vitamin D boosts cancer treatment

Vitamin D could make pancreatic cancer treatment more effective, by reprogramming cells that bolster tumour growth.

Pancreatic cancer is particularly deadly, partly because of cells called pancreatic stellate cells, which foster an environment that favours the growth of tumours and resists chemotherapy. Ronald Evans and Michael Downes of the Salk Institute in La Jolla, California, and their colleagues found that the vitamin D receptor is expressed in human pancreatic tumours. Activation of the receptor markedly altered

gene expression in pancreatic stellate cells, shifting them to a quiescent state in which they could not support tumours as well.

As a result, treating mice bearing pancreatic tumours with a vitamin D analogue and chemotherapy slowed tumour growth and increased survival compared with chemotherapy alone.

*Cell* 159, 80–93 (2014)

## ECOLOGY

### Dingo destruction okay for prey

Efforts to control Australia's dingo populations to protect livestock may not be having negative effects on other prey species.

Some studies have suggested



## PALAEONTOLOGY

### Amphibian regrew limbs long ago

The oldest evidence for limb regeneration has been found in fossils of a 300-million-year-old amphibian.

Salamanders can regrow entire lost limbs. Usually, the regrowths are indistinguishable from those that they replace, but in some cases they have distinctive abnormalities such as fused or missing digits. Nadia Fröbisch and her colleagues at the Natural History

Museum — Leibniz Institute for Evolution and Biodiversity Science in Berlin found similar abnormalities in fossils of *Micromelerpeton credneri* (pictured), a distant relative of modern amphibians.

This is the first fossil evidence for limb regeneration, and suggests that this ability originated in an ancient amphibian ancestor. *Proc. R. Soc. B* 281, 20141550 (2014)

CAROLA RADKE/MPFL

ALAIN HERZOG/EPFL

that controlling populations of top predators, such as the dingo (*Canis lupus dingo*), can indirectly cause declines in some prey species further down the food chain. Benjamin Allen at the University of Queensland in Gatton, Australia, and his colleagues laid poisoned bait for dingoes at several large study sites across the country. They found that prey populations in areas where dingoes were killed were similar to, or greater than, those in areas with no culling. Over the long-term, prey population sizes fluctuated independently of predator control levels.

This may be because the amount of dingo culling was not high enough to affect the animal's populations, the authors say, concluding that current dingo control practices probably do not need to be changed.

*Front. Zool.* 11, 56 (2014)

## ARCHAEOLOGY

## Stone tools not out of Africa

An advanced method of making stone tools did not spread out of Africa in a single wave as once thought, but evolved independently among different groups of early humans in Eurasia and Africa.

Stone-tool-making technology changed 400,000 to 200,000 years ago from a process whereby tools were made by chipping off and discarding flakes to shape a rock, to a more complex technique whereby the rock is first shaped (**pictured left**) in order to flake off pieces (**pictured, right**) for later use. Daniel Adler of the University of Connecticut in Storrs and his colleagues analysed artefacts, from a 325,000-year-old

archaeological site in Armenia, that were made by both methods and report that the objects were from the same archaeological layer.

The finding is the earliest evidence of the simultaneous use of the older 'bifacial' and the more complex 'Levallois' technologies outside of Africa, and suggests that the latter did not suddenly replace the former, the authors argue.

*Science* 345, 1609–1613 (2014)

## PALAEOCLIMATE

## Winds favoured Pacific exploration

Polynesians took advantage of an unusual shift in climate and tradewind direction about a 1,000 years ago to sail downwind towards New Zealand and other islands.

Ian Goodwin at Macquarie University in Sydney, Australia, and his colleagues reconstructed Pacific sea-level pressure and wind patterns during a period 700–1,200 years ago when certain Polynesian islands and New Zealand were colonized, and when the global climate shifted. They found that these climate changes resulted in altered wind patterns that allowed Polynesians to easily sail to the East Polynesian islands, New Zealand and Easter Island without having to travel against the wind.

The finding contradicts earlier assumptions that these voyagers needed to sail upwind to reach their destinations.

*Proc. Natl Acad. Sci. USA*  
<http://dx.doi.org/10.1073/pnas.1408918111> (2014)

## ASTROPHYSICS

## Space ripples could pump up stars

Gravitational waves could energize and brighten stars — possibly providing indirect evidence for the weak ripples in space time that are thought to

# SOCIAL SELECTION

Popular articles on social media

## Ig Nobel prizes provide fun fodder

In honour of the winners of this year's Ig Nobel Prizes, researchers on social media buzzed about holy images on toast, medical uses for bacon, the slipperiness of banana skins and other offbeat works of science.

The awards, presented by the *Annals of Improbable Research*, recognize quirky research papers that might otherwise have slipped into obscurity. Not many people were talking about 'Frictional coefficient under banana skin', for example, until it took home the physics prize. Shortly afterwards, Michael Lerner, a physicist at Earlham College in Richmond, Indiana, tweeted that the paper "is clearly showing up on one of my exams". Neil Cronin, a human-locomotion researcher at the University of Jyväskylä in Finland, tweeted: "Finding funding for muscle research: difficult. Finding funding for banana skin friction study: easy apparently."

*Tribol. Online* 7, 147–151 (2012)



Based on data from altmetric.com. Altmetric is supported by Macmillan Science and Education, which owns Nature Publishing Group.

➔ **NATURE.COM**  
For more on popular papers:  
[go.nature.com/kxf4e2](http://go.nature.com/kxf4e2)

be emitted by high-energy events such as exploding stars.

Barry McKernan at the City University of New York and his colleagues calculated the effect that gravitational waves would have on a star if the waves have frequencies matching those of the star's natural vibrations. They found that the star absorbs those waves, and if close to a powerful source such as merging black holes, it could heat up and brighten.

The study suggests that gravitational waves, which are difficult to detect, could interact more strongly with matter than previously thought.

*Mon. Not. R. Astron. Soc.* 445, L74–L78 (2014)

in zooplankton, suggesting early evolutionary origins for the hormone.

Melatonin is produced by many organisms, but its function in invertebrates has not been clear. Maria Antonietta Tosches, Detlev Arendt and their colleagues at the European Molecular Biology Laboratory in Heidelberg, Germany, studied larvae of the marine worm *Platynereis dumerilii*, which move up and down in the water at certain times of the day. The authors found that the larvae make melatonin in the brain and that production ramps up at night. This boosted neuronal activity, which resulted in less swimming, allowing the larvae to drift downward.

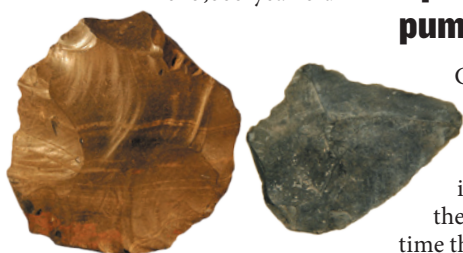
Melatonin evolved in early animals to coordinate their behaviour with the time of day, the authors propose. *Cell* 159, 46–57 (2014)

## PHYSIOLOGY

## Ancient roots of daily rhythm

The hormone that regulates sleep and other circadian processes in vertebrates also controls night-time behaviour

➔ **NATURE.COM**  
For the latest research published by Nature visit:  
[www.nature.com/latestresearch](http://www.nature.com/latestresearch)





# SEVEN DAYS

The news in brief

## POLICY

### Marine protection

US President Barack Obama has vastly expanded a marine reserve in the central Pacific Ocean, making it one of the largest in the world. On 25 September, Obama increased the Pacific Remote Islands Marine National Monument — which surrounds a group of small islands to the south and west of Hawaii — by more than 1 million square kilometres to 1.3 million. Former President George W. Bush created the reserve in 2009. See [go.nature.com/gpr5or](http://go.nature.com/gpr5or) for more.

### Wandering wheat

The US Department of Agriculture announced on 26 September that it had finished investigating an instance of escaped, experimental genetically modified wheat — and had begun looking into another such event. The agency said that it had exhausted all leads without determining the origin of herbicide-resistant wheat discovered on an Oregon farm last year (see *Nature* **499**, 262–263; 2013). It is now investigating the discovery of a different transgenic wheat strain at a research site in Montana. No transgenic wheat is approved for sale in the United States.

### Pathogen policy

The US government issued new rules on 24 September regulating a set of 15 pathogens and toxins related to ‘dual-use research of concern’ — life-science research that could be used for nefarious purposes. Researchers who receive government funding must report any work that they do with these agents to their institutions, which will assess potential hazards. In parallel with this, the federal



FREDERIC STEVENS/GETTY

## Large-scale ivory seizures on the rise

The worldwide trade in illegal ivory seemed to reach record levels between 2011 and 2013, driven by strong consumer demand in Asia, according to TRAFFIC, a wildlife-trade monitoring network based in Cambridge, UK. In a report published last week, the group found a rising trend in the frequency of seizures involving 500 kilograms of ivory or more since 2008. Elephant strongholds are collapsing:

for example, Tanzania's Selous Game Reserve once held more than 100,000 elephants; the population still exceeded 70,000 in 2007, but fell to just 13,000 in late 2013. TRAFFIC also reported a surge in rhinoceros poaching since 2008, after a lull from the early 1990s. In South Africa, rhino losses to poaching have grown every year since 2008, reaching 1,004 animals last year — around 5% of the 2012 population.

government will work with institutions to mitigate biosecurity risks. See [go.nature.com/h1sosc](http://go.nature.com/h1sosc) for more.

## FUNDING

### Mental-health grant

The US National Institute of Mental Health has funded a US\$16-million, 4-year study of the genetic basis of schizophrenia and bipolar disorder. Researchers at the University of Southern California in Los Angeles, the University of Michigan in Ann Arbor and the Broad Institute in Cambridge, Massachusetts, will sequence the whole genomes of at least 10,000 people, split between

those with schizophrenia, those with bipolar disorder, and healthy controls. In July, the Broad Institute also received a \$650-million donation to expand research into psychiatric conditions (see *Nature* **511**, 393; 2014).

### Ebola funding

On 25 September, the World Bank pledged to nearly double its commitment to fighting the Ebola outbreak in West Africa, to US\$400 million. The outbreak, which has claimed more than 3,000 lives, could have potentially disastrous economic and public-health impacts on the region (see page 15). On 23 September, the Wellcome

Trust biomedical-research charity in London announced a £3.2-million (US\$5.2-million) grant to accelerate clinical trials for Ebola therapies at existing treatment centres.

### School status push

On 26 September, Japan's education ministry announced that it would boost funding for 37 institutions designated as ‘super global universities’ to elevate their international research status. The institutions — including the University of Tokyo and Kyoto University — will hold the title for ten years, and will receive either ¥172 million (US\$1.6 million) or ¥420 million each year.

Universities will make their own plans for the money, but are expected to attract international faculty members and students.

## Gender progress

The London-based charity Equality Challenge Unit on 25 September announced its latest round of awards recognizing UK research institutions and departments with strong gender-equality policies. The University of Cambridge became only the fifth institution to win the silver Athena SWAN award, given to institutions that show effective action against previously identified challenges to gender equality. The UK National Institute of Health Research has suggested that from 2016 it will only shortlist medical schools for funding if they have received the silver award.

## EVENTS

## Volcanic eruption

Mount Ontake in central Japan erupted on 27 September at 11:53 a.m., spraying ash and debris on the surrounding region (**pictured**), as well as on hundreds of unsuspecting hikers on its slopes. As *Nature* went to press, at least 36 people were feared dead. Global-navigation stations and monitors of surface deformation related to rising magma did not detect any



unusual activity leading up to the eruption. See [go.nature.com/wpwymr](http://go.nature.com/wpwymr) for more.

## Flying on solar

Two Swiss pilots have planned the first round-the-world aeroplane flight powered only by solar energy. Last July, André Borschberg and Bertrand Piccard, co-founders of the Solar Impulse project, piloted the first all-solar plane flight across the United States. On 25 September, the team announced plans to circumnavigate the globe next year in a more-advanced solar aircraft. The fuel-less journey, which will begin and end in Abu Dhabi, is expected to take about ten legs between March and August 2015.

## Mars-club member

India has become the first Asian nation to get a craft into Mars orbit. Its Mangalyaan probe arrived at the red planet on 24 September, three days after NASA's MAVEN mission. Only the United States, the former Soviet

Union and the European Space Agency have previously sent successful missions to Mars. See [go.nature.com/ynzqsq](http://go.nature.com/ynzqsq) for more.

## Support for climate

The United Nations climate summit held in New York City on 23 September produced few firm pledges but generated international enthusiasm for climate-change policy. World leaders are scheduled to meet at the UN climate talks in December 2015 in Paris, where they are expected to discuss the successor to the 1997 Kyoto Protocol. See [go.nature.com/hs8ljh](http://go.nature.com/hs8ljh) for more.

## BUSINESS

## Chemicals buy-out

German chemicals firm Merck has agreed to pay US\$17 billion to buy Sigma-Aldrich, a company headquartered in St Louis, Missouri, that makes and sells more than 230,000 chemicals and biomolecules to researchers worldwide. Merck, headquartered in Darmstadt, said that the acquisition, announced last week, would make it a leading player in the global life-sciences industry.

## RESEARCH

## Super responders

On 24 September, the US National Cancer Institute launched a clinical trial to study

## COMING UP

### 6–8 OCTOBER

The Nobel prizes are announced for physiology or medicine, physics and chemistry.

### 9–10 OCTOBER

Physicists gather at the US National Institute of Standards and Technology in Gaithersburg, Maryland, to plan experiments that could determine a more reliable value for 'big G', the universal gravitational constant. [go.nature.com/wzya7c](http://go.nature.com/wzya7c)

people who respond unusually well to cancer therapies. The Exceptional Responders Initiative will deposit genomic data from 100 such people in a database, and will aim to use their biological and clinical data to identify others who may benefit from the same treatments. Advocates say that the approach could yield information about important genes and molecular pathways; critics counter that small sample sizes make it difficult to draw firm conclusions from such studies.

## Kepler back to work

NASA's Kepler spacecraft has returned its first batch of science data since resuming its hunt for planets outside the Solar System in June, the mission team announced on 23 September. The data include observations of more than 12,000 stars, as well as galaxies, to be scanned for supernovae or signs of black holes. The craft, hobbled since May 2013 by the failure of two of its reaction wheels, now steadies itself by balancing its frame against the oncoming solar wind. The team expects Kepler's fuel to last until at least late 2017.

➔ [NATURE.COM](http://NATURE.COM)

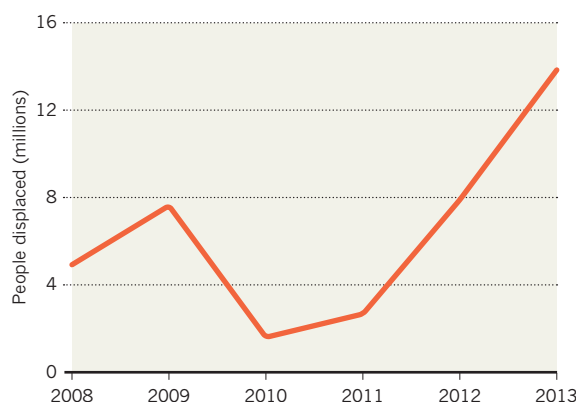
For daily news updates see: [www.nature.com/news](http://www.nature.com/news)

## TREND WATCH

Natural catastrophes displaced nearly 22 million people last year, with the largest displacements affecting populous Asian nations, according to a report by the Internal Displacement Monitoring Centre in Geneva, Switzerland. Storms forced 14.2 million people from their homes, including 13.8 million in Asia (see chart). In November 2013, 4.1 million people in the Philippines were displaced by Typhoon Haiyan — one of the largest typhoons ever recorded.

## RISING STORMS

Recent storm-related disasters have displaced increasing numbers of people across Asia.





# NEWS IN FOCUS

**BIOPIRACY** Protocol will stop exploitation — and create red tape **p.14**

**BOTANY** Forensic chemistry to stop South Africa's plant thieves **p.17**

**ASTRONOMY** Telescope data bounty sparks access debate **p.18**

**ASTRONOMY** Physicists debate future of Argentina's cosmic-ray observatory **p.20**

CONRAD BORNHMAN/GALLO/GETTY



The genomes of ill newborns can be sequenced in less than 24 hours to give clinicians a rapid diagnosis.

## GENOMICS

# Fast sequencing saves newborns

*Rapid analysis of infant genomes is aiding diagnosis and treatment of inexplicably ill babies.*

BY SARA REARDON

By two months of age, the boy was near death. He had spent his entire short life in the neonatal intensive care unit (NICU) at Children's Mercy Hospital in Kansas City, Missouri, while physicians tried to work out the cause of his abnormalities. When his liver failed in April 2013, the medical staff warned his parents that the outlook was grim.

Then geneticist Stephen Kingsmore and his team at Children's Mercy took on the case. Within three days, they had sequenced the genomes of the baby and his parents, and identified a rare mutation that was common to the child and both of his parents. The mutation turned out to be linked to a disease in which an overactive immune system damages the liver and spleen. Armed with a diagnosis, the baby's physicians put him on drugs to lower his immune response. The boy is now at home

and healthy. Had physicians sent his DNA off for a conventional genomic test, the diagnosis could have taken more than a month — by which time he would probably have died.

The boy is one of 44 sick infants whose genomes Kingsmore's group has sequenced using a process that can provide a diagnosis in as little as 24 hours. In 28 of these cases, the researchers have been able to diagnose the baby's condition. And in about half of these, they have been able to recommend changes in treatment, Kingsmore reported on 19 September at the Genomics of Common Diseases meeting in Potomac, Maryland. On 6 October, his group will kick off a larger project to sequence hundreds of babies' genomes. It will be the first of four newborn-sequencing studies that each received multimillion-dollar grants from the US National Institutes of Health (NIH) in September 2013. The studies will address both the feasibility and the ethics of a process that could soon become standard for inexplicably ill newborns.

Over the next five years, Kingsmore's group will sequence the genomes of 500 sick babies from the Children's Mercy Hospital NICU and compare the infants' clinical outcomes with those of 500 NICU babies who are diagnosed using conventional genetic and metabolic tests. The researchers will assess whether rapid sequencing allows babies to avoid unnecessary tests and unhelpful treatments, and whether it helps parents to make decisions about care when the child is diagnosed as having a fatal disease. Even when an infant does die, Kingsmore says, a genome sequence and diagnosis can provide closure to parents and give more information about the genetic conditions they carry.

Kingsmore calls the rapid sequencing technique a 'factory' approach, in which four or five specialists each perform one step of the process — from the blood draw to the final diagnosis — as quickly as possible. The group collects DNA from both of the parents and the baby to quickly identify mutations in the child's genome, then sequences the DNA and uses custom software to target specific parts of the genome on the basis of their symptoms. After making a gene-based diagnosis and delivering relevant information to the baby's physician, the group stores the sequence data anonymously in a secure database for use in future studies.

Misha Angrist, a genomic-policy expert ▶



► at Duke University in Durham, North Carolina, says that although the 24-hour genome process is impressive, it is not clear whether genomic sequencing of newborns will soon become standard practice. Many questions remain about who will pay for sequencing, who should have access to the data and how far clinicians should go in extracting genome information that is unrelated to the disease at hand. Then there is the question of how informative the process is. “I think it’s really important that we do these experiments so that we start to see what that yield is,” Angrist says.

So far, only the Kansas City team has been cleared to begin trials, thanks to a waiver from the US Food and Drug Administration (FDA) that allows sequencing of very ill babies. Normally, a test must be experimentally proven before being used to diagnose patients. “These are very pioneering studies,” Kingsmore says. “I think that everybody is keen to see whether this is the start of a new approach at FDA, and whether it will happen in the future with similar studies.”

The other NIH-funded teams are awaiting approval from the FDA or from internal ethics-review boards. In Boston, Massachusetts, a group led by physicians Alan Beggs of Boston Children’s Hospital and Robert Green of the Brigham and Women’s Hospital is planning a study of 240 healthy babies and 240 from NICUs. The team will randomly sequence the exome — the protein-encoding portions of the genome — for half of each group of infants to determine whether those data alone can improve children’s health. Exome sequencing is cheaper, albeit less comprehensive, than whole-genome sequencing.

A third team, led by geneticists Cynthia Powell and Jonathan Berg of the University of North Carolina in Chapel Hill, plans to sequence the genomes of 400 babies with known genetic diseases, such as cystic fibrosis, to see whether they can glean extra information about the disorders. And medical geneticist Robert Nussbaum’s group at the University of California in San Francisco will sequence exomes from 1,400 bloodspots, previously collected from infants at birth, to determine whether this information is useful for diagnosis.

Each team includes ethicists who will grapple with questions such as disclosing information that is unrelated to the diagnosis. “People are sensitive about the power of information in genomics and rightly so,” Green says. Those concerns are magnified when they involve children. ■

**“Everybody is keen to see whether this is the start of a new approach.”**



A shaman in Ecuador gathers plants to make ayahuasca, which was at the centre of a biopiracy row.

## POLICY

# Biopiracy ban stirs red-tape fears

*Critics worry Nagoya Protocol will hamper disease monitoring.*

BY DANIEL CRESSEY

A major international agreement is coming into force to combat ‘biopiracy’ — profiting from biological products while failing to compensate the community from which they originate. The Nagoya Protocol aims to ensure that developing nations benefit when their plants, animals or microbes are used by foreign scientists.

But some researchers fear that the agreement will stymie vital activities, such as conservation or monitoring and treating infectious diseases.

The protocol takes effect on 12 October, four years after it was signed in Nagoya, Japan. Its 92 signatories include Brazil, Japan and the European Union. Notably absent are China and the United States, although researchers in those countries will have to comply with the laws of nations where they collect samples.

Part of the United Nations Convention on Biological Diversity (CBD), the protocol has the stated purpose of ensuring “fair and equitable sharing of benefits arising out of the utilization of genetic resources”, which covers all organisms. Researchers must already obtain permits to collect samples from certain countries, but the protocol means that they will have

to enter into ‘access and benefit sharing’ (ABS) arrangements. These set out who might profit — and how — from the organisms being used, and stipulate how to distribute the benefits fairly, for example through co-authorship of publications, or sharing profits from products such as drugs, vaccines or crops.

Several high-profile cases underscore the need for such rules, says Braulio de Souza Dias, executive secretary of the CBD secretariat. In a case often cited as a victory against biopiracy, a European patent on an antifungal agent derived from neem, an evergreen tree native to India, was revoked in 2000 after a long legal battle, on the grounds that Indian farmers had used the fungicide for decades. Other controversies have involved a US patent on the use of turmeric in wound healing, which was withdrawn, and one on ayahuasca — a hallucinogenic tea made from Amazonian plants — which has now expired.

The importance of the issue also became apparent in 2007, when Indonesia balked at sharing samples from people infected with avian influenza with the World Health Organization, on the grounds that the nation would not benefit from any resulting papers or patents. Indeed, scientists working abroad stand to gain from the

WADE DAVIS/GETTY

protocol, says Dias, because it will build trust between them and local people, which could lead to better access to organisms. In the past, “no one trusted anyone”, he says. The protocol could also help countries to access treatments that are developed using disease samples taken from their own people.

But although scientists understand the need for ABS agreements, many worry that they will have destructive consequences.

The protocol has the potential to hamper disease monitoring, according to the London-based biomedical research charity the Wellcome Trust. Red tape could make it harder to quickly share samples across borders, which in turn could cripple efforts to monitor drug resistance in malaria, for example, or outbreaks of *Escherichia coli*. “There need to be equitable arrangements for sharing benefits, but it is absolutely critical that policy-makers ensure they do not hinder these international partnerships that are so vital to protect global public health,” says David Carr, a policy adviser at the Wellcome Trust.

The new rules will also present challenges for synthetic biologists, who combine genetic code from many different organisms to create

drugs or sensors. This could require dozens of ABS arrangements for a single product, says Tim Fell, chief executive of Synthace, a biotechnology company in London. Such bureaucracy could push European companies to countries — particularly the United States — that are not signatories, he adds.

**“If I compare two sequences, is that utilization? I don’t know.”**

International research collaborations may face a bureaucratic challenge if their members operate under different laws, says the London-based BioIndustry Association.

There is also uncertainty about the protocol’s reach, particularly for genetic sequences. A possible interpretation of the rules is that anyone who uses sequence data would have to complete ABS paperwork. Christopher Lyal, who studies weevils at London’s Natural History Museum, helps to run a CBD website that provides advice about the protocol. Even he is unsure of how it will affect him: “If I compare two sequences to reach a conclusion on identification, is that utilization? I don’t know.”

The BioIndustry Association also says

that the threat of criminal charges for non-compliance — the UK government is considering jail terms of up to two years — could have a chilling effect on research.

Some researchers think that the protocol could even hurt the countries it is intended to help. Kazuo Watanabe, director of the Gene Research Center at the University of Tsukuba in Japan, fears that red tape surrounding access to and exchange of specimens will hinder field studies in disciplines such as taxonomy and ecology. This, in turn, will make it harder to help conservation efforts.

Dias acknowledges the potential problems, but says that people will have to deal with them: “There will be a cost for a transition phase, yes, but it should be for the better.”

Elisa Morgera, who specializes in global environmental law at the University of Edinburgh, UK, agrees. There may be uncertainty in the short term, with “difficult negotiations and possible missteps”, she says, but the protocol offers a way to rebuild trust. “Those genuinely interested in the long-term viability and reputation of bio-based research and innovation would be well advised to constructively contribute to this process,” she says. ■

## INFECTIOUS DISEASE

# Ebola obstructs malaria control

*Outbreak is shutting down prevention and treatment programmes in West Africa.*

BY ERIKA CHECK HAYDEN

As the Ebola death toll spirals into the thousands in West Africa, the outbreak could have a spillover effect on the region’s deadliest disease. The outbreak has virtually shut down malaria control efforts in Liberia, Guinea and Sierra Leone, raising fears that cases of the mosquito-borne illness may start rising — if they haven’t already.

So far, at least 3,000 people are estimated to have died of Ebola in Guinea, Sierra Leone and Liberia in the current outbreak, although World Health Organization (WHO) staff acknowledge that official figures vastly underestimate the total. By contrast, malaria killed more than

6,300 people in those countries in 2012, most of them young children. Overall, malaria deaths have fallen by about 30% in Africa since 2000 thanks to national programmes supported by international funding agencies such as the Global Fund to Fight AIDS, Tuberculosis and Malaria, the US Agency for International Development and the WHO’s Roll Back Malaria initiative. The schemes distribute free bed nets to protect sleeping children from mosquitoes, train health workers to find malaria cases and offer tests and treatment at no charge to patients.

But the Ebola outbreak has brought those efforts to a standstill in the three affected countries. “Nobody is doing a thing,” says Thomas Teuscher, acting executive director of the Roll

Back Malaria Partnership, based in Geneva, Switzerland.

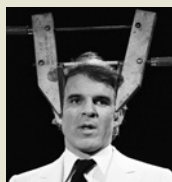
He says that malaria drugs are sitting in government warehouses, especially in Liberia and in Guinea, where medical supply trucks have been attacked by people angry with the government’s handling of the Ebola outbreak. Liberia had planned a national campaign to distribute bed nets this year, but Teuscher says that it may be difficult to launch that now.

Routine health care has collapsed during the outbreak, because both patients and providers have shunned clinics for fear of infection. As a result, tens of thousands of people could die from treatable causes, says Estrella Lasry, a tropical-medicine specialist for medical



**MORE  
ONLINE**

## TOP STORY



Physicists manage to squeeze quantum data for the first time  
[go.nature.com/nb7vqv](http://go.nature.com/nb7vqv)

## MORE NEWS

- Earth has water older than the Sun [go.nature.com/ft3jwn](http://go.nature.com/ft3jwn)
- ‘Weak nuclear force’ explains life’s asymmetry [go.nature.com/lqhngq](http://go.nature.com/lqhngq)
- Nobel laureates want Iranian physicist freed from Teheran prison [go.nature.com/1htceb](http://go.nature.com/1htceb)

## EXPLAINER



Were warnings of Japan eruption missed?  
[go.nature.com/xntje5](http://go.nature.com/xntje5)



► charity Médecins Sans Frontières (also known as Doctors Without Borders) in New York. Those include complications of childbirth; trauma and other acute conditions requiring surgery; and causes such as diarrhoeal disease, respiratory viruses and especially malaria. With proper treatment malaria can usually be cured completely, but if left untreated it can develop into a severe form that is often fatal.

“It’s a disaster in all ways possible,” says Lasry. “The public-health impact will be huge.”

As of August, the WHO had not seen a year-on-year increase in people with malarial symptoms reporting to clinics in Guinea, the only Ebola-affected country where such data

are available. In fact, malaria deaths in Guinean clinics decreased for the first half of this year compared with 2013. But that is not necessarily good news, says Teuscher. It could mean that the illiest people have been staying away from clinics, scared off by the Ebola outbreak, and their deaths have gone unnoted.

Furthermore, the symptoms of malaria mimic Ebola, so many people who might have malaria are avoiding clinics for fear of learning the worst, says Alice Johnson, a nurse and clinical mentor for Last Mile Health, an organization in Boston, Massachusetts, that trains health workers in rural Liberia.

Ebola’s impact on malaria programmes is likely to linger long after the outbreak ends. In Guinea, for instance, authorities bury Ebola victims with their bed nets to prevent the spread of infection; this has raised suspicion that the nets have some inherent connection to Ebola.

And health workers are afraid to do blood tests to confirm malaria because Ebola is spread by blood and other bodily fluids. That could lead to people who do not have malaria being given antimalarial medication, which can contribute to the development of drug resistance in the parasite that causes the disease.

It is important to get malaria control programmes back on track, says Teuscher, in part because they could help to fight Ebola.

For instance, in Sierra Leone about 2,000 community health workers have been trained to go into villages to find and treat malaria. They could also be trained to detect Ebola and help infected people to get care, he says.

“Potentially, we have an army of people available in these countries who have experience delivering malaria treatments,” says Teuscher. “They’re still there; they just need to be helped to do a good job.” ■

## CLIMATE SCIENCE

# Tibetan plateau gets wired up for monsoon prediction

*Largest and highest plateau in the world has outsized impact on climate.*

BY JANE QIU IN LHASA

The gigantic, remote Tibetan plateau is being flooded with sensors in an unprecedented attempt to understand its influence on climate — especially the Asian monsoons, which caused deadly flooding in India and Pakistan in September. The US\$49-million Chinese effort could help to predict extreme weather — both in Asia and as far afield as North America — and give scientists a steer on how climate change affects these events.

Sitting at an average height of around 4,000 metres above sea level, the plateau protrudes into the middle of the troposphere, where most weather events originate. As the biggest and highest plateau in the world, it disturbs this part of the atmosphere like no other structure on Earth. But there are little data on the impact that this has on climate.

In central and western Tibet, where weather observations are particularly lacking, researchers jointly funded by the China Meteorological Administration and the National Natural Science Foundation of China began, in August, to place temperature and moisture detectors in the soil and to erect 32-metre-high towers laden with sensors that measure cloud properties. In recent weeks, the team has begun deploying sensors mounted on weather



The Tibetan plateau, often called the third pole, will be monitored by balloons, drones and ground sensors.

balloons and unmanned aerial vehicles.

Such sensors will eventually monitor a vast swathe of the plateau’s ground and air — across diverse landscapes such as desert, grassland, forest and farmland. “The data should help determine the extent to which different types of land surface heat up the overlying air, and how this might vary in response to factors

such as snow cover and vegetation changes,” says Wu Guoxiong, an atmospheric scientist at the Institute of Atmospheric Physics of the Chinese Academy of Sciences (CAS) in Beijing and a principal investigator of the project.

Scientists agree that Tibet plays a key part in the climate system, but many of the details are a mystery. The plateau’s remoteness, altitude

WOLFGANG KAHLER/LIGHTROCKET VIA GETTY



and harsh conditions — it is often called the third pole because it hosts the world's third-largest stock of ice — mean that even basic weather stations are few. Satellite data are also plagued by large errors owing to lack of calibration from ground observations.

"Climate models have the greatest uncertainties in Tibet and the Himalayas, and are especially weak at simulating monsoons," says Xu Xiangde, an atmospheric scientist at the Chinese Academy of Meteorological Sciences in Beijing and investigator on the project. This dearth of information about the plateau, acknowledged by the Intergovernmental Panel on Climate Change, affects scientists' ability to predict how the climate is changing, and the consequences for people living in vulnerable regions.

The plateau's altitude means that it receives more sunlight and so gets hotter than land at sea level. And because land absorbs more solar radiation than air, the plateau acts like a giant heating plate. This heat pumps air upwards, which disperses in the upper troposphere, giving the plateau an outsized influence over atmospheric circulation, and thus climate. The heating effect also intensifies monsoons, which arise as a result of a temperature difference between land and the oceans that sets up a pressure gradient in the atmosphere. In 2008, Wu reported that the surface heating of the plateau had been weakening since the 1980s (A. Duan & G. Wu *J. Clim.* **21**, 3149–3164; 2008), consistent with a weakening in the strength of Asian monsoons. But monsoons seem to be getting stronger again, and occurring in places where they were previously rare, says Klaus Fraedrich, an atmospheric scientist at the University of Hamburg in Germany.

In early September, a deadly flood caused by a monsoon hit border regions between India and Pakistan that are normally dry, killing hundreds and affecting millions more. If the Chinese project can help to explain why monsoons are changing, it "could help instigate early evacuation plans and save many lives", says Fraedrich.

The project could have yet broader effects. A team led by Hai Lin, an atmospheric scientist at Environment Canada in Quebec, found that the greater the snow cover in Tibet, the warmer the winter in Canada (H. Lin & Z. Wu *J. Clim.* **24**, 2801–2813; 2011). The latest initiative could confirm Lin's suspicion that increased snow cover causes the plateau to reflect more sunlight, reducing its heating capability and strengthening a pressure system that causes warmer-than-usual winters in North America. Ma Yaoming, an atmospheric scientist at the CAS Institute of Tibetan Plateau Research in Beijing, says that combined with data on glaciers, permafrost, rivers and lakes, the project will contribute to a better picture of Asia's entire water cycle. ■



The illegal trade in South Africa's cycads is threatening to push the endangered species to extinction.

#### BOTANY

## Forensic chemistry could stop plant thieves

*Scientists hope to save rare cycads using isotope analysis.*

BY LINDA NORDLING IN CAPE TOWN

Scarred earth meets visitors at the Kirstenbosch National Botanical Garden where some of South Africa's rarest plants once stood. In August, 24 of the garden's cycads were stolen, probably to be sold on the black market as landscaping ornaments. As with the country's emblematic rhinos, time is running out for the plants. But scientists hope that a forensic method that is also used to track ivory might help to deter plant poachers.

South Africa's endemic cycads rank among the most endangered plants in the world. Of the country's 38 species, 3 are extinct in the wild and 12 are critically endangered. The plants grow slowly and can live for hundreds of years. Their striking looks and rarity make them prized collectors' items, with individual plants fetching tens of thousands of US dollars.

This profitability fuels illegal poaching, which has proved hard to stop even though it carries a ten-year prison sentence. Microchip tags embedded in the plants have been spotted by thieves armed with X-ray machines, and gouged out. And it is not feasible to treat every plant in a collection — let alone in the wild — with a more successful method that sprays plants with microdot paint containing identification tags that are too small to be seen with the naked eye.

A team led by plant

scientist Adam West of the University of Cape Town hopes that chemistry can help. The forensic method used by the team depends on the fact that the relative abundances of a chemical element's isotopes vary naturally from place to place. As organisms grow, they incorporate these isotope signatures, providing a trace of where they have lived. Stable-isotope analysis has helped to identify the origins of smuggled ivory, counterfeit money and drugs.

When West's team used the method to compare the isotope signatures of cycads that they knew had been relocated with those of wild plants that had never been moved, they found that it was possible to identify the relocated plants. Their results, to be published in the November issue of the *Journal of Forensic Sciences*, suggest that the method can reveal a plant relocation that happened decades ago. "If you got your cycad from the wild 30 years ago, we can still tell," says West. The team is now testing suspect plants that were flagged in police raids, to see whether the isotope signatures are consistent with the owner's story or with a wild origin. West hopes that the ability to read a plant's history might deter illegal dealers.

It is "an elegant piece of work", says Jason Sampson, curator of the Manie van der Schijff Botanical Garden in Pretoria. But he says that more also needs to be done to sate the demand for rare cycads, for instance by accelerating breeding programmes. ■

➔ **NATURE.COM**

For a longer version of this story, see:

[go.nature.com/jbrn6g](http://go.nature.com/jbrn6g)

# Data bounty spurs debate

*Small institutions fear exclusion from Large Synoptic Survey Telescope.*

BY MARK ZASTROW

**N**ow under construction atop a mountain in northern Chile, the 8.36-metre Large Synoptic Survey Telescope (LSST) will sweep the entire southern sky every three nights when it starts operating in 2022 — creating a wealth of data that will be available to all US astronomers and dozens of international partners. It promises to be a democratizing force and to usher in a new era of survey astronomy.

But that promise could go unrealized without the proper infrastructure, astronomers warn. Without access to the tools and facilities needed to analyse the huge data set and to do follow-up observations, many astronomers could be locked out of the bounty. Especially vulnerable are researchers and students at small and minority-serving institutions, which often find it hard to secure telescope time.

The US National Science Foundation (NSF), which is footing the telescope's US\$473-million construction bill, has commissioned a National Research Council (NRC) panel to formulate a strategy that maximizes the scientific return of the LSST. It is a complicated problem, says the panel's chair, Debra Elmegreen. To help it decide, the panel has asked astronomers to provide input by 6 October on how they intend to use the LSST and what support they would need to be able to do so. The panel's report is due early next year.

A big part of the facility's appeal is that it will detect unexpected events such as supernovae or stars being swallowed by black holes — but exploring details such as their composition and temperature will require access to other ground-based telescopes. Large US research universities typically have private access to such telescopes, but small ones tend to rely on public instruments, which are under threat from budget cuts. In 2012, a panel recommended that the NSF divest itself of several facilities, which would halve the number of nights open to visiting observers. The agency plans to follow the recommendation, but has been delayed by a budget stalemate in the US Congress.

## QUESTION OF CAPACITY

Another common concern is that analysing big data sets requires correspondingly large computing resources. The LSST will collect so much data (30 terabytes per night) that few small institutions will have the capability to analyse the information directly. "You're not going to copy the whole LSST data set,"



Smaller telescopes will be needed to investigate events spotted by the Large Synoptic Survey Telescope.

says Joshua Pepper, an astronomer at Lehigh University in Pennsylvania. "Even a subset is beyond the range of a professor and an office desktop."

One solution is to create an online portal that would let astronomers mine the database remotely. There is a smaller-scale precedent:

**"There are hungry minds and students and young scientists who are eager and ready to participate."**

the Sloan Digital Sky Survey (SDSS), which uses a 2.5-metre telescope at the Apache Point Observatory in New Mexico. Its portal enables anyone to view and filter the telescope's

output; as a result, the data have been used in more than 5,800 publications that have been cited 245,000 times.

LSST director Steven Kahn notes that there have always been plans for an online portal. But the flow of data will be so massive that even basic processing is an enormous job, says Keivan Stassun, an astronomer with joint appointments at Vanderbilt University and Fisk University, both in Nashville, Tennessee, who chairs the SDSS executive committee. The LSST will collect more data in three nights than the entire SDSS catalogue, so Stassun worries that despite its best intentions, the LSST could find itself lacking resources. "That's not a criticism of LSST; it's a statement of capacity," he says.

The make-up of the NRC panel has also raised eyebrows. The only member from a small institution is Elmegreen, an astronomer

at Vassar College in Poughkeepsie, New York. Because the panel's main remit is to maximize the LSST's scientific return, she considers its primary mission to be ensuring wide availability of data. "I'm a little bit conflicted," she says, "because I'd like to make sure that everyone has access to telescopes. But the big push today is to make sure that people have access to data."

LSST director Kahn dismisses concerns about the composition of the panel. "The LSST committee is completely committed to the idea of open access and serving the whole community."

Still, Stassun expresses disappointment that there are no representatives on the panel from minority-serving institutions such as Fisk, which was founded to serve African American students and has no guaranteed access to private telescopes. "There are hungry minds and students and young scientists who are eager and ready to participate in the enterprise who have traditionally been excluded," says Stassun. "And yet again, there's not a seat at the table." ■

## CORRECTION

The News story 'Seed-patent case in Supreme Court' (*Nature* **494**, 289–290; 2013) implied that Monsanto patented a method for engineering transgenic crops to produce sterile seeds before 1999. Although it began negotiations in 1998 to acquire the firm that filed the patents, the deal only completed in 2007. Monsanto never commercialized such crops.

DAVID NUNUK/GETTY



# TO CATCH A COSMIC RAY

When ultra-high-energy cosmic rays arrive from interstellar space, they strike air molecules and produce a cascade of lower-energy particles.

The Pierre Auger Observatory in Argentina has spent almost ten years looking for the source of ultra-high-energy cosmic rays — but to no avail. Now the observatory faces an uncertain future.

BY KATIA MOSKVITCH

The tank looks oddly out of place here on the windy Pampas of western Argentina. Surrounded by yellow grass and spiky thorn bushes, the chest-high plastic cylinder could be some kind of storage container — were it not for the bird-spattered solar panels and antennas on top.

More tanks can be seen in the distance, illuminated by a crimson Sun dropping behind the far-off Andes. “Some locals think that the tanks influence the weather: they make it rain or snow, or make a dry season,” says Anselmo Francisco Jake, the farmer who owns this stretch of land. “But I know they don’t. I know they catch cosmic rays.”

Jake is right. There are 1,600 of these tanks, spaced over a 3,000-square-kilometre expanse that could fit all of Luxembourg with room to spare. Together they comprise the Pierre Auger Observatory: a US\$53-million experiment to reveal the mysterious origins of ultra-high-energy cosmic rays, the most energetic subatomic particles known to exist.

But for all its size, the array has fallen short. After almost ten years of hunting, it has observed dozens of ultra-high-energy cosmic rays, but has not managed to solve the mystery of where they come from. As a detector, “the device worked twice as well as we expected”, says project co-founder James Cronin, a retired astrophysicist at the University of Chicago in Illinois. But the particles seem to be coming from all over the sky, with too little clustering for researchers to pinpoint the sources. “It’s up to nature with experiments like this one,” he says.

Now, the Auger team is putting its hopes on a proposed upgrade that might settle the question by improving Auger’s resolution considerably. Five designs are being evaluated internally by a committee of Auger physicists, who are expected to present their final selection to the

array’s many funding agencies in November. The trouble is, there is a sixth option, too. “In the worst-case scenario, and I don’t want to think about it, we may get shut down,” says Auger’s deputy project manager, physicist Ingo Allekotte.

An upgrade would require an investment of roughly \$15 million, and some argue that the money would be put to better use elsewhere. “Although it was worth building Auger, it was a gamble that unfortunately didn’t yield much new understanding,” says Eric Adelberger, a physicist at the University of Washington in Seattle. “Cosmic-ray physics has delivered very few surprises and progress is terribly slow. Maybe it is time to move on.”

*“It’s up to nature with experiments like this one.”*

That would be a blow to science — and to Argentina, say Auger’s supporters. These flagship projects do more than just conduct research, says Pablo Mininni, head of the physics department at the University of Buenos Aires. They also raise awareness of physics and draw young people into the field. “Such a big project deserves some continuity,” he says.

Physicists have known for more than a century that Earth is continually bombarded by charged particles from space — many of which have energies that are astonishing

even by particle-physics standards. It is not uncommon for cosmic rays to have hundreds or thousands of times the 7 trillion electron volts ( $10^{12}$  eV) soon to be achieved by the most powerful human-made particle accelerator, the Large Hadron Collider (LHC) near Geneva in Switzerland.

Most of these particles are now thought to be protons and other light nuclei originating far outside the Solar System, probably in cataclysmic stellar explosions known as supernovas. But on very rare occasions, cosmic rays have hit Earth’s atmosphere at energies of  $10^{18}$  eV or more. The most energetic example on record — the ‘Oh-My-God particle’ detected<sup>1</sup> on 15 October 1991 in the skies above Utah — had  $3 \times 10^{20}$  eV, about 40 million times that of the LHC. And therein lies a mystery: calculations suggest that the expanding shock wave of a supernova detonation cannot accelerate charged particles beyond about  $10^{17}$  eV. No one knows what physical process could accelerate particles to higher energies — or even what those particles might be (see *Nature* 448, 8–9; 2007).

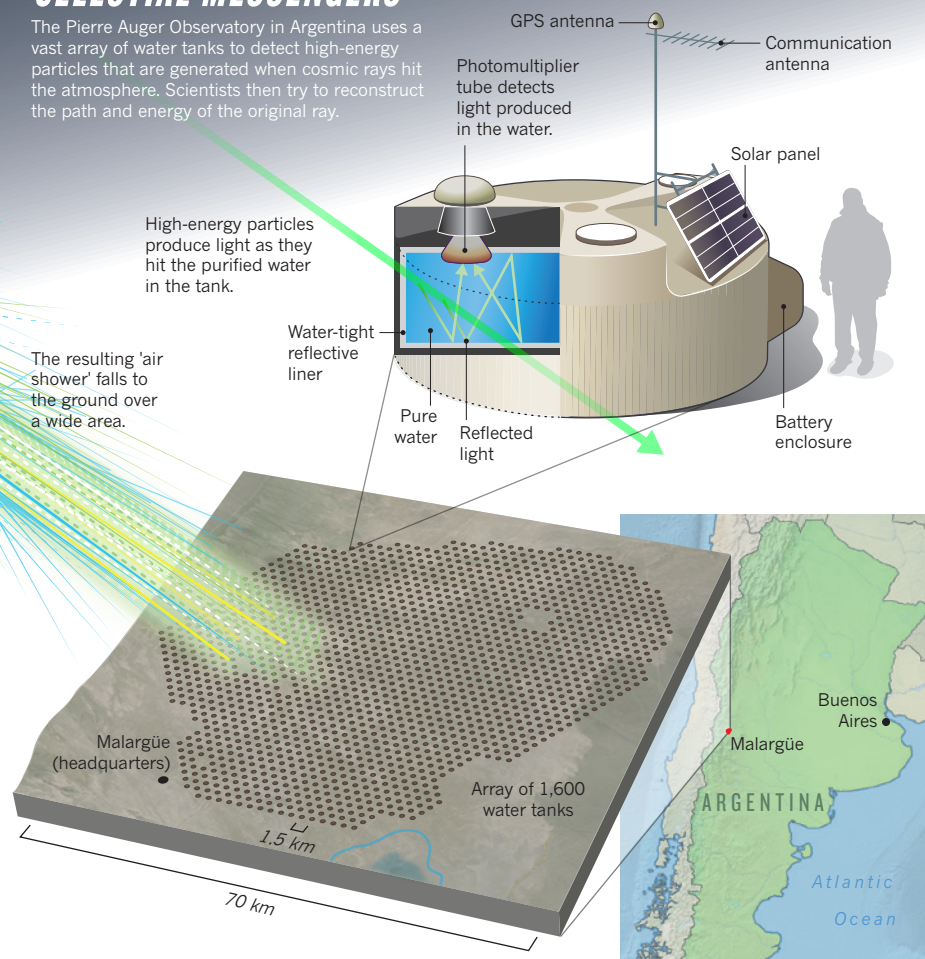
## RULE-BREAKERS

In 1992, Cronin, who shared the 1980 Nobel Prize in Physics for his work on particle interactions, decided to find out. He, Alan Watson of the University of Leeds, UK, and Murat Boratav of Pierre and Marie Curie University in Paris, set out to build an observatory that — they hoped — could detect enough ultra-high-energy cosmic rays to answer those questions.

Their sprawling, 1,600-detector design reflected two fundamental facts about their quarry. The first is that the rays are exceedingly rare. Although their low-energy cousins come in at roughly a few particles per square centimetre per second, the rates dive

## CELESTIAL MESSENGERS

The Pierre Auger Observatory in Argentina uses a vast array of water tanks to detect high-energy particles that are generated when cosmic rays hit the atmosphere. Scientists then try to reconstruct the path and energy of the original ray.



by the idea of his country hosting an international science project that he promised to support it with the equivalent of US\$10 million in Argentinian pesos. The province of Mendoza, where the site is located, agreed to contribute another \$5 million.

This largesse would prove to be a mixed blessing: in 2001, just as construction was getting under way, Argentina experienced its biggest economic crisis and government default in history. The peso instantly lost two-thirds of its value, leaving the researchers to scramble for funding from other sources to keep construction on schedule. It was one of Auger's biggest setbacks, says Cronin. Another came in 2010, when US funding agencies declined the researchers' request to build a sister observatory in Colorado, which would have allowed them to look for the ultra-high-energy cosmic-ray sources across the entire sky instead of just the Southern Hemisphere.

### AIR SHOWERS

Still, the first 154 detectors of the Auger observatory were able to start collecting data on 1 January 2004, with the rest of the detectors being deployed in stages until the array was completed in 2008. Each of the plastic tanks is filled with 12,000 litres of purified water, which produces a streak of light when an air-shower particle passes through, and is lined with phototubes that can measure that light. The tank's antennas transmit the data to the observatory's headquarters in Malargüe, where they are sent out for analysis to some 350 researchers around the world.

Their first decade of data-taking has yielded a number of provocative results, including hints that many of the highest-energy rays are actually heavy nuclei such as that of iron, instead of the much more common protons<sup>2</sup>. "It was a surprising result that nobody had thought about," says Auger spokesman Karl-Heinz Kampert, a physicist at the University of Wuppertal in Germany. And if true, it might have something important to say about the mysterious acceleration mechanism — although no one is quite sure what. It also threatened to undermine Auger's central quest: heavy nuclei tend to be more strongly deflected by intergalactic magnetic fields than protons are, and that could randomize their direction and make it impossible to trace the rays back to their sources.

That concern seemed to have been put to rest in 2007. Working with three-and-a-half years of data gleaned from 27 rays, Auger researchers reported that the rays seemed to preferentially come from points in the sky occupied by supermassive black holes in nearby galaxies<sup>3</sup>. The implication was that the particles were being accelerated to their ultra-high energies by some mechanism associated with the giant black holes. The announcement generated a media frenzy, with reporters claiming that the mystery of the origin of cosmic rays had been solved at last.

precipitously as the energy increases. Above  $10^{20}$  eV, the cosmic-ray flux is less than one particle per square kilometre per century. So the more detectors the researchers could deploy, the better their chances would be of catching one.

The second fact is that 'primary' cosmic rays — those that are coming in from interstellar space — never reach the ground. Instead, they smash into an air molecule high in the atmosphere, producing a blast of photons, electrons, positrons, muons and other collision products that then slam into other air molecules. The result is an 'air shower': a cascade of lower-energy particles that collectively follow along the track of the original cosmic ray. And that calls for detectors over a very wide area, in the hope that the devices could register enough of the air-shower particles as they hit the ground to reconstruct the energy and direction of the original particle (see 'Celestial messengers'). To help in the reconstruction, the physicists also planned to surround the site with four clusters of fluorescence telescopes to scan the skies over the array, mapping the faint streaks of blue and ultraviolet light that the air-shower particles produce as they rip through the atmosphere.

Naming their observatory after Pierre Auger, the French physicist who discovered air showers in 1938, the three scientists started going from

country to country knocking on doors. They gathered a cadre of high-level physicists from around the world who wanted to join them. And those physicists, in turn, used their connections to get funding from their own governments. In short order, the United States agreed to help, as did Italy, Germany, France, Argentina and several other countries.

At the same time, the Auger team was looking at potential sites in South Africa, Australia and South America — places that met their need for lots and lots of empty, flat land with clear skies above. Nelson Mandela dearly wanted the observatory to be based in South Africa. But the Auger team judged that the nation did not have a strong-enough community of physicists to support the project.

The Australian site had a different drawback: it was on land controlled by the military, so collaborators from certain countries might not be able to work there.

So in November 1995, Cronin, Watson and Boratav announced that the observatory would be built in Pampa Amarilla, a plain some 1,400 metres above sea level. Except for Malargüe, a mining town of 23,000 people just to the southwest, the site was as empty as the Auger team could wish. Better still, Argentina's then-president, Carlos Menem, was so excited





The Pierre Auger Observatory's detectors sit incongruously in western Argentina's yellow prairie.

But it had not. As the years went on and as the data accumulated, the correlations got weaker and weaker. Eventually, the researchers had to admit that they could not unambiguously identify any sources<sup>4</sup>. Maybe those random intergalactic fields were muddying the results after all. Auger “should have been more careful” before publishing the 2007 paper, says Avi Loeb, an astrophysicist at Harvard University in Cambridge, Massachusetts.

The Auger physicists contend that it would have made no sense to wait. “We gave the statistical significance of what we observed, so scientists know how to ponder the results,” says team member Esteban Roulet, a physicist at the Balseiro Institute in San Carlos de Bariloche, Argentina. “I think it is important that the community gets the information we can gather in this way.”

#### MASSIVE UPGRADE

Nonetheless, the mystery remains unsolved — an impasse that the Auger team wants to end with the hoped-for upgrade. The basic strategy is to get a better measure of each primary cosmic ray's mass and thus distinguish the relatively undeflected protons from the heavier particles, says Auger team member Alberto Etchegoyen, a physicist working at Argentina's National Atomic Energy Commission in Buenos Aires. “If nature is kind enough to us,” he says, and if there are enough protons among the ultra-high-energy cosmic rays to get adequate statistics, “we'll be able to find the sources”.

Currently, the mass is measured by Auger's fluorescence telescopes, which watch how each air shower expands and deposits its energy as it descends through the atmosphere. But the telescopes can operate only on clear, moonless nights, which cuts down on their observing

time. So instead, the team wants to look within the showers to count muons: short-lived particles that behave like heavy electrons. Because the muons in air showers tend to be produced most copiously in collisions of heavier cosmic-ray particles, knowing their abundance should tell the Auger physicists whether the incoming primaries are protons or heavy nuclei.

The five upgrade proposals represent five different ways of identifying muons, but all are

*“We have managed  
to grow a whole  
new generation of  
experimentalists.”*

based on the fact that the muons tend to penetrate farther into the water tank than other particles. Each scheme requires a different combination of new electronics, new detectors and internal modifications for all 1,600 tanks — hence the \$15-million cost of the upgrade. Supporters argue that the investment is worthwhile, not least because the array currently has little chance of ever getting statistics good enough to identify the sources, yet still costs \$1.7 million a year to run.

But the muon-detection schemes have yet to be proved in the field, and the selection

committee could still decide that the upgrade is not worth it — and perhaps even that Auger should be shut down. “This is a serious question,” says Kampert.

Cronin insists that it is much too soon to give up. Auger is exploratory, he says. “I don't know how much we'll learn from it. But you don't learn anything if you don't do something.”

Besides, says Allekotte, scrapping Auger would be depriving Argentina of a project that has greatly boosted the country's scientific capacity — not least by providing an incentive for young people to pursue physics. Tiny Malargüe now has a university for first- and second-year undergraduate students where many Auger engineers and scientists teach part-time. “One girl who started in 2012 was at first interested in maths, but as she learned more and more about the observatory and cosmic rays, she decided to switch to physics,” says Marcos Cerda, an Auger engineer and a physics lecturer at the university. “She's now in her third year, doing a physics major at the University of Mendoza.”

In addition, says Etchegoyen, there are many Argentinian students among the roughly 360 who have already earned their PhDs doing research at Auger, or are working towards one there. And now, he says, “two out of Auger's five upgrade proposals — design, prototype construction, everything — would be made in Argentina. That would've been impossible at the beginning of Auger. We have managed to grow a whole new generation of experimentalists linked to international big physics.”

Thanks to the observatory, “Argentina appeared on the map of global science”, says the country's science minister Lino Barañao. He points to the Deep Space Antenna 3 radio dish that the European Space Agency installed about 30 kilometres south of Malargüe to support space missions such as Mars Express, Herschel and Planck. And he points to the Large Latin American Millimetre Array, a radio telescope being built in the north of Argentina in collaboration with Brazil. The presence of Auger influenced the decisions to base both these projects in Argentina.

So if the Auger upgrade does go ahead, Argentina hopes to gain even more expertise, and add more capacity, says Barañao. And even if it doesn't, at least it's left a legacy. “We're associated with producing soya beans, beef and wine, but many countries can do that,” he says. “Now we're also associated with world-class astrophysics.” ■

*Katia Moskvitch is a science writer in London and an International Development Research Centre fellow at Nature.*

1. Bird, D. J. et al. *Astrophys. J.* **424**, 491 (1994).
2. The Pierre Auger Collaboration *Phys. Rev. Lett.* **104**, 091101 (2010).
3. The Pierre Auger Collaboration *Science* **318**, 938–943 (2007).
4. The Pierre Auger Collaboration *Astropart. Phys.* **34**, 314–326 (2010).





# EXTREME LIVING

After humans arrived in South America, they quickly spread into some of its most remote corners.

BY BARBARA FRASER

From the mouth of a cave high in the Andes, Kurt Rademaker surveys the plateau below. At an altitude of 4,500 metres, there are no trees in sight, just beige soil dotted with tufts of dry grass, green cushion plants and a few clusters of vicuñas and other camel relatives grazing near a stream.

The landscape looks bleak, but Rademaker views it through the eyes of the people who built a fire in the rock shelter, named Cuncacha, about 12,400 years ago. These hunter-gatherers were some of the earliest known residents of South America and they chose to live at this extreme altitude — higher than any Ice Age encampment found thus far in the New World. Despite the thin air and sub-freezing night-time temperatures, this plain would have seemed a hospitable neighbourhood to those people, says Rademaker, an archaeologist at the University of Maine in Orono.

“The basin has fresh water, camelids, stone for toolmaking, combustible fuel for fires and rock shelters for living in,” he says. “Basically, everything you need to live is here. This is one of the richest basins I’ve seen, and it probably was then, too.”

Rademaker is one of a growing number of young archaeologists investigating how

hunter-gatherers first colonized South America at the close of the Pleistocene epoch, when the last Ice Age was waning. Casting aside old dogmas, these researchers are finding that people arrived significantly earlier than previously believed, and adapted rapidly to environments from the arid western coastline to the Amazon jungle and the frosty heights of the Andes.

By teaming up with geologists, climate scientists and other researchers, archaeologists are gaining a clearer picture of what the ancient environments were like and how people migrated across the landscape — clues that are leading them to other ancient occupation sites.

## HIDDEN ANCESTRY

“The archaeology that’s being done in South America is becoming more scientific with the development of new methodologies, and there’s a level of collegiality developing among younger researchers,” says Rademaker. “We’re all really excited about the new developments that are coming faster and faster.” But researchers are racing against time as South American countries rapidly expand mining, road building and other activities that threaten to obliterate evidence from promising sites.

For decades, a fractious attitude prevailed over research on the earliest people in the Americas. One of the most acrimonious disputes concerned a site in southern Chile called Monte Verde, which Tom Dillehay, an anthropologist now at Vanderbilt University in Nashville, Tennessee, excavated in the 1970s and 1980s. He found evidence of human occupation<sup>1</sup> that he dated to about 14,500 years ago. Dillehay’s conclusions regarding Monte Verde put him in direct conflict with the accepted wisdom among leading archaeologists that people from Siberia did not spread across North America and venture south before around 13,000 years ago. That is the age of the Clovis culture, a group of big-game hunters who used distinctive spear points that are found littered across the United States. The Clovis people were thought to be the pioneers in North America, and many archaeologists there dismissed Dillehay’s claim that Monte Verde was older.

But antagonism has faded over the past six years, as convincing evidence of pre-Clovis sites has emerged in North America (see *Nature* **485**, 30–32; 2012). Meanwhile, South American archaeologists, who were never as sceptical as their northern colleagues, have

BARBARA FRASER





“These Palaeo-Indians were able to live in one of the most extreme environments on Earth, at the end of an ice age, and they seem to have done so quite successfully.”

found more sites dated between 14,000 and 12,000 years ago, indicating that hunter-gatherers had spread through South America before and during the rise of the Clovis culture in the north.

Now that researchers have moved beyond that debate, they are making greater headway in studying when people reached South America and what they did when they got there.

Rademaker's finds in the Andes are helping to answer those questions — and pose new ones. His journey began 150 kilometres away from the Andes cave, on Peru's arid coast at Quebrada Jaguay, where Daniel Sandweiss, an anthropologist at the University of Maine and Rademaker's graduate adviser, was excavating a site that dated to the end of the last Ice Age, between 13,000 and 11,000 years ago. Sandweiss had uncovered the remains of seafood



**Above:** Christopher Miller (left) and Rademaker survey sites in the Pucuncho Basin in August.  
**Left:** Kurt Rademaker explores the Cuncaicha rock shelter in the Andes.

meals, as well as flakes of obsidian produced as people chipped at the glassy mineral to make stone tools<sup>2</sup>. There are no obsidian deposits along that coastline, so the material must have come from formations high in the Andes.

Rademaker travelled into the mountains and found a large outcrop of the obsidian known as Alca<sup>3</sup> at Mount Condorsayana in 2004. Over the next three years, he studied the obsidian deposits and evidence of past glaciation in the area with geologist Gordon Bromley of the University of Maine.

Those field trips gave Rademaker his first glimpse of the Pucuncho Basin, an alpine wetland with a stream, numerous vicuñas, llamas and alpacas, and a ready supply of cushion plants, which the researchers discovered are rich in resin and can burn easily. The basin was also littered with points and shards left by early toolmakers. Hiking down the stream, he glanced up the hill to his left and saw a yawning gap — the Cuncaicha rock shelter, which he began excavating in 2007.

“This is the first time we've found a site this old in the high Andes,” Rademaker says. On a day in August, he wraps a bandana over his mouth and nose and shovels dirt into buckets to fill in an excavation pit that is no longer needed. As he works, his shirt sleeve pulls up, revealing a glimpse of meticulously detailed hominin skulls tattooed up his right arm — from *Australopithecus afarensis* near his wrist to *Homo sapiens* on his shoulder. This late in the field season, his field trousers are frayed and he has had to bind his left hiking boot with several strata of duct tape.

A chilly breeze whips across the Pucuncho plateau as some of Rademaker's companions struggle with the thin air. As well as cautioning his team members to prepare for the cold, Rademaker ensures that they acclimate gradually to the lack of oxygen.

Even while battling the extremes, the team has gathered evidence contradicting the conventional wisdom that the mountains were too high, cold and inhospitable for early human habitation. Bromley's data show that at the end of the last Ice Age, glaciers were mainly

confined to some alpine valleys, and Pucuncho and other areas were not glaciated. Palaeoclimate data indicate that the environment was probably wetter then, so there might have been more plants and animals available for the early residents, says Rademaker.

“These Palaeo-Indians were able to live in one of the most extreme environments on Earth, at the end of an ice age, and they seem to have done so quite successfully,” he says. “This tells us that Palaeo-Indians were capable of living just about anywhere.”

There are large numbers of animal bones, mainly from deer and vicuñas, in the earliest layers of sediment in the Cuncaicha rock shelter, showing that the inhabitants found abundant game on the plateau. And some of the tools were made of stone not available in the area, indicating that residents of the cave either travelled outside the region or exchanged materials with other groups that did. Some tools show traces of plant starch, which the researchers hope to analyse to work out what the cave-dwellers ate, and whether they domesticated tubers or other plants.

The researchers have also found a fragment from a human skull at the site. It has not yielded DNA and its age is uncertain, but it hints that the cave could contain early human remains, says Rademaker.

## TOOL TRADE

Farther south, César Méndez has followed similar clues in his search for late-Pleistocene sites along the Chilean coast. Beginning in 2004, Méndez, an anthropologist at the University of Chile in Santiago, and his colleagues excavated an ancient encampment, which they dated to around 13,000 years ago<sup>4</sup>.

Some of the stone tools at the site, called Quebrada Santa Julia, were made of translucent quartz that is not found in coastal deposits. Like Rademaker, Méndez mapped potential paths towards known quartz deposits inland. Sampling along those routes, his team found an outcrop of translucent quartz at a site where people had lived and quarried between 12,600 and 11,400 years ago. The similarity with Quebrada Santa Julia in terms of age and tool-making

BARBARA FRASER



techniques suggests that the coastal tools came from these mountain outcrops.

“What we’re seeing is that 12,000 years ago or more, these groups already had networks, knew the landscape and moved between the coast and the interior,” says Méndez.

Sites such as Quebrada Jaguay and Quebrada Santa Julia suggest that some early hunter-gatherers in South America might have travelled along the coast, taking advantage of the fish, shellfish, animals and plants found in wetlands and near river deltas, says Dillehay. He is finding more evidence beneath Huaca Prieta, a 32-metre-high mound on the coast of northern Peru (see ‘Conquering a continent’).

The mound was first excavated in the 1940s, but Dillehay dug deeper and uncovered traces of Ice Age settlements in 2010. Radiocarbon dating indicates<sup>5</sup> that humans had lived there as much as 14,200 years ago, when the area was surrounded by wetlands.

### COASTAL DRIFT

If early people did migrate along the coast, some of the best evidence has probably been swallowed up by the ocean. At the end of the Pleistocene, melting ice sheets caused sea levels to rise by 70 metres, which would have flooded much of the former coastline. That effect would have been greatest in some regions of eastern South America, where the land is relatively flat and the ocean migrated well inland.

At the border between Uruguay and Argentina, for example, archaeologists suspect that ancient people might have hunted and camped on a broad delta that formerly existed at the mouth of the Uruguay River. But any such sites would have been drowned when the sea advanced by more than 120 kilometres, says Rafael Suárez, an archaeologist at the University of the Republic in Montevideo.

Suárez has looked for clues upriver, and has dated several residential sites to between 12,900 and 10,200 years ago. Some tools found at a site called Pay Paso are made of translucent agate, which apparently came from quarries near the border with Brazil about 150 kilometres away. And other tools from Uruguay have been found 500 kilometres to the south in Argentina’s Buenos Aires province<sup>6</sup>, says Nora Flegenheimer, an archaeologist with the National Scientific and Technical Research Council (CONICET) in Necochea, Argentina. Such finds point to widespread trade or travel routes in eastern South America.

Some archaeologists wonder whether early residents of the continent might even have crossed the Andes. Bolivian archaeologist José Capriles of the University of Tarapacá in Arica, Chile, has raised that possibility after studying 12,800-year-old artefacts at Cueva Bautista, a rock shelter 3,930 metres above sea level in southwestern Bolivia. He notes that a similarly aged site exists at the same latitude in Chile on the western slope of the Andes. Future research could explore tools found at both sites to see



whether people migrated from one side to the other or established trading routes.

But some of the best evidence for Pleistocene humans in South America may disappear soon, owing to rapid expansion in industrial-scale agriculture, road building and other forms of development. Those human threats come on top of the natural ones — wind erosion and changing watercourses — that constantly alter landscapes.

Suárez and his team had to call the navy to evacuate them from a site in Uruguay last December, when floodwaters rose dangerously in the lake behind a nearby hydroelectric dam. A proposed dam could also flood sites in the Ocoña River valley in Peru, which Rademaker thinks could have been an early route from the coast to the Andes.

In the highlands, the rapid expansion of mining can be both a bane and a blessing. Archaeologists discovered Bolivia’s Cueva Bautista site during a survey for a road leading to a mine. But open-pit mines threaten many other sites, says Capriles.

Archaeological surveys must be carried out before development and infrastructure projects can go ahead, but the people who perform such studies do not always recognize the subtle signs of ancient human occupation, the researchers say. And even if the surveys do turn up important archaeological evidence, developing countries are often reluctant to let the past stand in the way of the future.

“I’ve never seen such destruction as you get in Peru,” says Dillehay. He has witnessed bulldozers ravage sites and landowners destroy evidence to avoid delaying construction work.

There are no signs yet of such activity reaching Rademaker’s survey site in the high Peruvian Andes. Over the past decade, he and his colleagues have extensively explored

the region on foot in an effort to determine whether the inhabitants of the Cuncaicha rock shelter traded for their exotic tools and whether they lived there year-round. The answers may lie in undiscovered occupation sites between the cave and the coast, so Rademaker is exploring likely avenues, mapping the routes that would have required the least energy expenditure while providing access to water and food.

The researchers have backpacked along dozens of streams and rivers, sometimes clambering up steep cliffs to avoid flash floods, always with an eye out for gashes in the rock face that signal a potential shelter. Early inhabitants probably would have explored the new landscape in the same way with the same targets in mind.

Rademaker surveyed four rock shelters this year but all of them were inhabited too recently — only 4,000 to 6,000 years ago. Still, he is convinced that there are more late-Pleistocene sites in the Andes. Early inhabitants must have found other places like the Pucuncho Basin and the Cuncaicha rock shelter. They might have followed rivers that flow from the highlands to the coast. Or perhaps they trailed the herds of wild guanacos that still descend along spurs of the Andes nearly to the ocean shore.

Each field season dangles more possibilities before Rademaker’s team. “I went for a walk one night, found another confluence and found another cave,” he says. “It’s never-ending.” ■

**Barbara Fraser** is a writer in Lima, Peru.

1. Dillehay, T. D. *et al. Science* **320**, 784–786 (2008).
2. Sandweiss, D. H. *et al. Science* **281**, 1830–1832 (1998).
3. Rademaker, K. *et al. Geology* **41**, 779–782 (2013).
4. Méndez, C., Jackson, D., Seguel, R. & Nuevo Delaunay, A. *Curr. Res. Pleistocene* **27**, 19–21 (2010).
5. Dillehay, T. D. *et al. Quat. Res.* **77**, 418–423 (2012).
6. Flegenheimer, N., Bayón, C., Valente, M., Baeza, J. & Femenías, J. *Quat. Int.* **109–110**, 49–64 (2003).



# COMMENT

**CLIMATE** Replace 2°C warming goal with meaningful vital signs **p.30**

**INNOVATION** A chronicle of the teamwork that drove the digital revolution **p.32**

**EVOLUTION** Model shows it is a short step from squid to albatrosses **p.34**



**AGEING** Study optimism and resilience as well as physical decline **p.35**

CHLOE DEWE MATHEWS/PANOS



A boat rusts on the bed of the dried Aral Sea, more than 90% of which has vanished in the past 50 years.

## Curb vast water use in central Asia

Irrigation-intensive industries in former Soviet republics have sucked water bodies dry. **Olli Varis** calls for economic reform to ease environmental and social tensions.

Shipwrecks rusting in the desert have come to symbolize the environmental havoc that has befallen the Aral Sea, which straddles Kazakhstan and Uzbekistan. More than 90% of what was once the fourth-largest lake in the world has vanished in half a century<sup>1-3</sup>. The cracked shores are symptoms of the dramatic overuse of water in central Asia. Since the 1960s, 70% of Turkmenistan has become desert, and half of Uzbekistan's soil has become salty owing to

dust blown from the dry bed of the Aral Sea<sup>1</sup>.

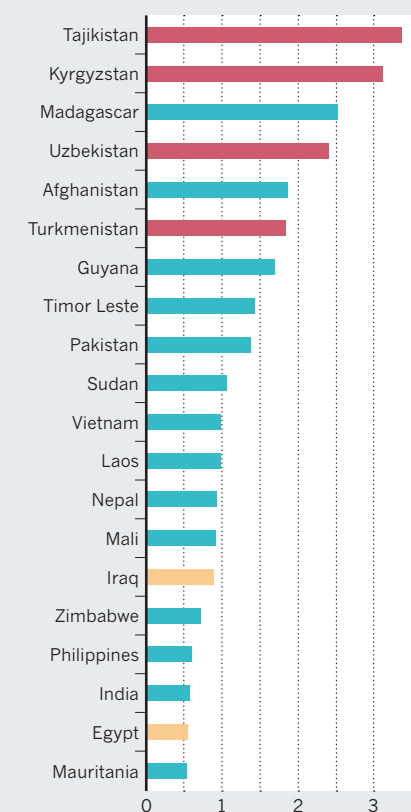
The republics of Uzbekistan, Tajikistan, Turkmenistan, Kyrgyzstan and Kazakhstan were developed as farming states to supply produce to the former Soviet Union<sup>1</sup>. Today, they are among the highest per capita users of water in the world — on average, each Turkmen consumes 4 times more water than a US citizen, and 13 times more than a Chinese one<sup>4</sup> (see 'Top 20 consumers'). More than 90% of the region's water use is irrigating thirsty

crops including cotton and wheat<sup>1,2</sup>.

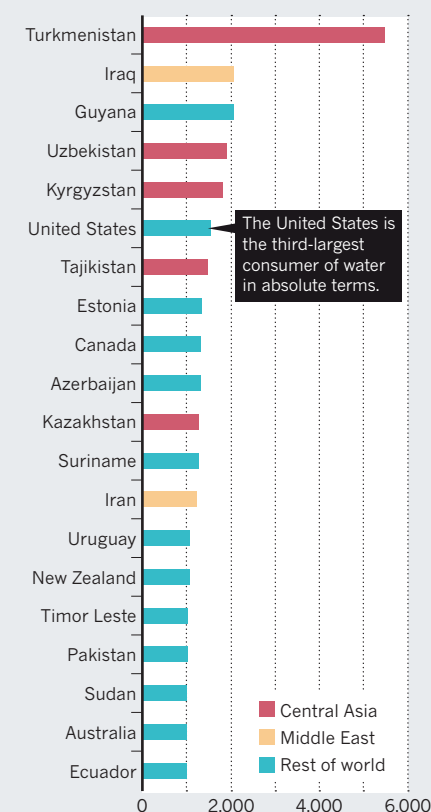
Decades of over-extraction have nearly sucked dry the Amu Darya and Syr Darya rivers that feed the Aral Sea. Local livelihoods that rely on livestock grazing, hunting and fishing have disappeared; ecosystems in the Aral's brackish waters, deltas, coasts, steppes and fertile river valleys have collapsed<sup>1</sup>. As water bodies have vanished, the local climate has become harsher: summers bring extreme heat and violent, salty dust storms; winters

## TOP 20 CONSUMERS

Central Asian republics use disproportionately large quantities of water relative to the size of their economies and populations. Most water goes to irrigate crops grown in poor-quality soils.

Water use (m<sup>3</sup>) per US\$ of GDP\*

\*Gross domestic product

Water use (m<sup>3</sup>) per capita

The United States is the third-largest consumer of water in absolute terms.

Central Asia  
Middle East  
Rest of world

► are more severely cold. The wind spreads salt and agrochemicals to farmlands hundreds of kilometres away, causing respiratory and gastroenterological diseases as well as anaemia, cancer and tuberculosis<sup>3,5</sup>.

Struggling to shake off the Soviet legacy of environmental and political crises and oligarchies, these republics are more rivals than neighbours. Because most of the region's water bodies — mainly the Syr Darya, Amu Darya and Zarafshon rivers — are shared, political tensions have grown around water access, drawing worrying parallels with similar crises in the Arab world.

The first step is to recognize that the origin of central Asia's water problems is in excessive water demand. Fixing the problem will mean developing regional industries that are less water intensive and more profitable than agriculture, by tapping human potential rather than natural resources. Unless the region's economy can be put on a more sustainable footing, the stability and security of central Asia is in danger.

## SHORTAGE MYTH

Two fallacies stymie debate about water in central Asia. The first is that the region is short of water. The landscape looks dry

and rivers run empty. Many analyses in the past few years<sup>3,5</sup> have thus recommended water-conservation measures, assuming that incremental policy changes are all that can be delivered. In fact, these countries have plenty of water relative to their populations.

***"In fact, these countries have plenty of water relative to their populations."***

The annual availabilities of fresh water per capita for the Amu Darya (2,087 cubic metres) and the Syr Darya (1,744 m<sup>3</sup>) river basins<sup>6</sup> are well above the United Nations definitions of water shortages<sup>7</sup>: 1,000 m<sup>3</sup> per capita constitutes a chronic shortage, and 1,700 m<sup>3</sup> a moderate shortage. By comparison, Denmark has 1,128 m<sup>3</sup> of water per capita, Germany 1,878 m<sup>3</sup> and the United Kingdom 2,465 m<sup>3</sup> (ref. 4).

The second fallacy is that the solution is agricultural. Most analysts propose that water should be used more efficiently on farms because it is wasted in growing low-return crops on dry lands unsuitable for agricultural use. Turkmenistan's dry climate and poor soils mean that producing a tonne of wheat takes 2,000–4,000 m<sup>3</sup> of irrigation water, whereas in nearby

northern Kazakhstan adequate rainfall and conditions mean that no irrigation is needed. Even as its land became parched, Turkmenistan's wheat yield increased nine-fold between 1992 and 2007.

But the big fish swims elsewhere: the agricultural share of gross domestic product (GDP) in central Asia has almost halved since the disintegration of the Soviet Union<sup>4</sup>. Instead, economic growth is dominated by the oil and gas industry and by urban expansion. Already, more than half of the region's population is urban and that proportion is rising.

Despite this, central Asian economies continue to focus on primary industries such as agriculture and the extraction of fossil fuels. The economic return on water is lower in central Asia than anywhere else on the planet. Turkmenistan uses nearly 3 times more water than India to produce one GDP dollar, 4 times more than Egypt, 14 times more than China and 43 times more than Spain<sup>4</sup>.

## RISING TENSION

The resulting problems are greater than just stagnant economies. Disputes (see 'Troubled waters') between nations have arisen around access to shared water bodies in the Fergana Valley in the Syr Darya river basin, in the Zarafshon river basin, and in Amu Darya — most notably concerning the Nurek dam and Turkmen–Uzbek rivalries on water appropriation.

These tensions are stoked by absurd projects such as the Golden Age Lake (Altyn Asyr) in the Karakum Desert<sup>8,9</sup>. Projected to cover almost half the area of the Great Salt Lake in Utah, the synthetic lake will be about six times its volume. Since 2000, Turkmenistan has been constructing it, claiming it will increase agricultural production and offer a "symbol of revival of the Turkmen land", as former president Saparmurat Niyazov (known as Turkmenbashi) put it<sup>9</sup>.

Water for the lake will be drawn from the Amu Darya river through two canals, which are being cut across about 3,200 km of desert<sup>8,9</sup>. Although it is unclear whether that much water can ever be sourced from the river, it is obvious that downstream, Uzbekistan will not accept those diversions and is ready to defend its water share with arms if necessary. The already serious soil-salinization problems of Turkmenistan and Uzbekistan will be greatly worsened if the project is completed.

Like most other parts of the former Soviet Union, central Asian states suffer authoritarian rule and political fragility. Soaring unemployment is leading to a mass emigration of educated people. Current figures estimate that up to one-third of working-age Tajiks are employed abroad. Ethnic, political



and religious diversity and difficulties with boundary demarcation fuel nationalism. Internal hostilities, as in the Caucasus, Moldova and eastern Ukraine, are a threat. A full-scale regional conflict, regardless of the rise of radical religious groups, is not out of the question.

Central Asia's water crisis echoes that in the Middle East and North Africa, where political, economic and environmental issues are also intertwined. In Arab countries such as Syria, Yemen and Tunisia, water is scarce and used for low-value purposes, generating little income or investment<sup>10</sup>. Urban populations are fast-growing but ill-served by development policies focused on traditional rural and primary industries. Political and professional inertia makes change difficult.

Three main differences may make the situation in the former Soviet republics worse than in the Middle East. First, investments in the central Asian water sector are even less productive and more conflict-prone than in Arab countries. Second, water is more abundant in central Asia but environmental disasters have been more severe there than in Arab countries. Third, Arab cities absorb immigrants more successfully and grow faster than those in central Asia, where skilled workers tend to emigrate to countries outside the region, notably Russia.

The central Asian countries must find joint interests and competitive advantages to build a new regional economy, with wise water use at its heart. These countries could have a much more conscious role in world politics and in the global economy by looking at their complementary strengths and merging their markets.

### HUMAN POTENTIAL

The human resources of central Asia are relatively untapped. The republics have essentially full adult literacy and well over



90% of adults have secondary education<sup>8</sup>. The nations are in a favourable geographical position between diverse markets, including China, Russia, the Middle East and Europe.

Different national strengths should be exploited: Turkmenistan is rich in oil, Tajikistan and Kyrgyzstan in hydropower, for instance. Urban economies, services, manufacturing and knowledge-intensive industries should be boosted by governance reforms.

Realizing human potential would require policies to attract investments, maintain and enhance high standards of education, help industries to grow, and empower a bigger share of the population to contribute to political decision-making. Inertia may be the real bottleneck.

Experience from elsewhere abounds. Information and communication technology brings in more than one-quarter of India's export earnings; China, South

Korea, Vietnam and some other ex-Soviet states — notably Estonia — have also created knowledge-based industries almost from scratch. Such industries provide intellectually attractive, high-income jobs for the younger generation and put little strain on water resources and the environment.

International policy-makers and the water sector must refocus and look much more broadly at water's role in the region's political and economic development. That wider perspective should guide the next round of water-resources assessments, as well as top-level international policy meetings such as the 7th World Water Forum in Daegu, South Korea, in April 2015.

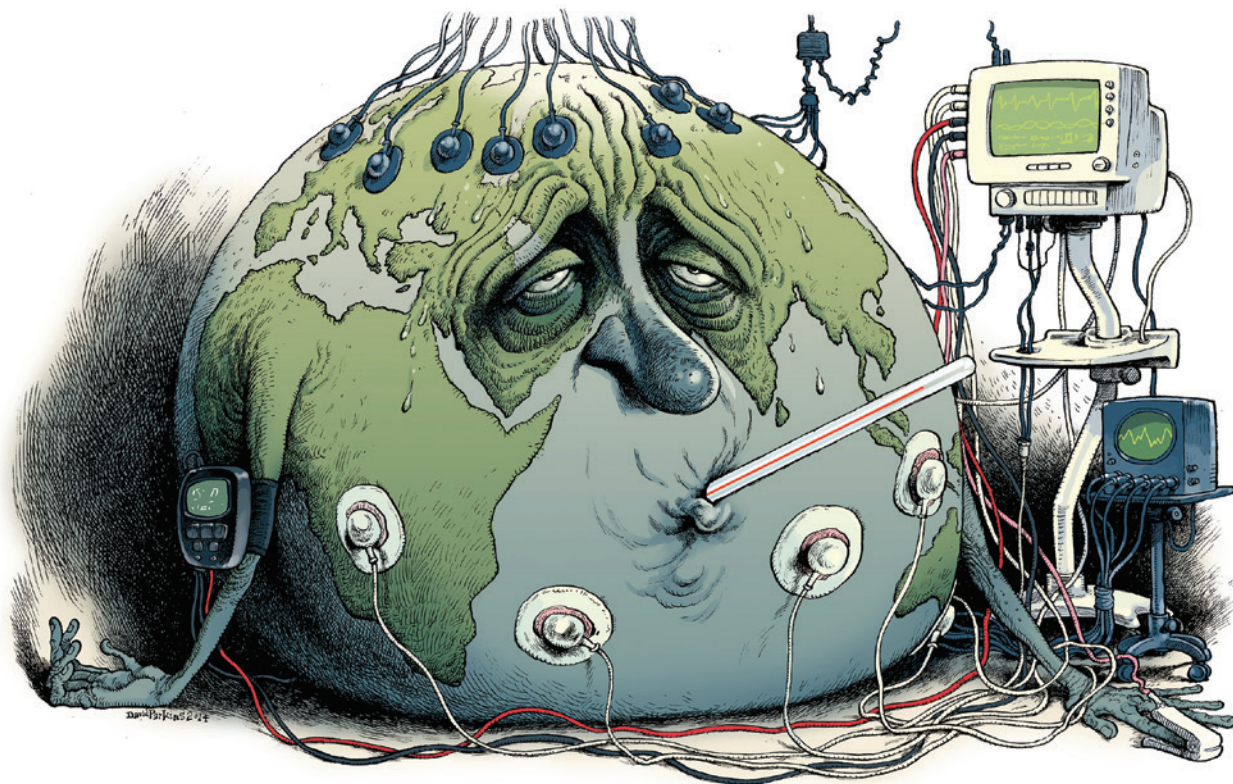
The alternative could be much worse: more iron wreckage on the drylands — this time of military origin. ■

**Olli Varis** is professor of water-resources management at Aalto University, Finland.  
e-mail: [olli.varis@aalto.fi](mailto:olli.varis@aalto.fi)

1. Dukhovny, V. A. & de Schutter, J. *Water in Central Asia: Past, Present, Future* (CRC, 2011).
2. United Nations Economic Commission for Europe *Strengthening Water Management and Transboundary Water Cooperation in Central Asia: The Role of UNECE Environmental Conventions* (UNECE, 2011).
3. Granit, J. et al. *Regional Water Intelligence Report Central Asia* (Stockholm International Water Institute, 2010).
4. World Bank *World Development Indicators* (World Bank, 2013).
5. United Nations Environmental Programme *Environment and Security in the Amu Darya Basin* (UNEP, 2011).
6. Porkka, M., Kumm, M., Siebert, S., & Flörke, M. *Int. J. Water Resour. Dev.* **28**, 455–476 (2012).
7. Falkenmark, M. *Phil. Trans. R. Soc. Lond. B* **352**, 929–936 (1997).
8. Zonn, I. S. & Kostianoy, A. G. (eds) *The Turkmen Lake Atyr and Water Resources in Turkmenistan* (Springer, 2013).
9. Stone, R. *Science* **320**, 1002–1005 (2008).
10. Varis, O. & Abu-Zeid, K. *Int. J. Water Resour. Dev.* **25**, 507–522 (2009).



Kazakhstan's capital Astana: rapid urban expansion will influence the region's water use.



# Ditch the 2°C warming goal

Average global temperature is not a good indicator of planetary health. Track a range of vital signs instead, urge **David G. Victor** and **Charles F. Kennel**.

For nearly a decade, international diplomacy has focused on stopping global warming at 2°C above pre-industrial levels. This goal — bold and easy to grasp — has been accepted uncritically and has proved influential.

The emissions-mitigation report of the Fifth Assessment of the Intergovernmental Panel on Climate Change (IPCC) is framed to address this aim, as is nearly every policy plan to reduce carbon emissions — from California's to that of the European Union (EU). This month, diplomatic talks will resume to prepare an agreement ahead of a major climate summit in Paris in 2015; again, a 2°C warming limit is the focus.

Bold simplicity must now face reality. Politically and scientifically, the 2°C goal is wrong-headed. Politically, it has allowed some governments to pretend that they are taking serious action to mitigate global warming, when in reality they have achieved almost nothing. Scientifically, there are better ways to measure the stress that humans are placing on the climate system than the growth of average global surface temperature — which has stalled since 1998 and is

poorly coupled to entities that governments and companies can control directly<sup>1</sup>.

Failure to set scientifically meaningful goals makes it hard for scientists and politicians to explain how big investments in climate protection will deliver tangible results. Some of the backlash from 'denialists' is partly rooted in policy-makers' obsession with global temperatures that do not actually move in lock-step with the real dangers of climate change.

New goals are needed. It is time to track an array of planetary vital signs — such as changes in the ocean heat content — that are better rooted in the scientific understanding of climate drivers and risks. Targets must also be set in terms of the many individual gases emitted by human activities and policies to mitigate those emissions.

## OWN GOAL

Actionable goals have proved difficult to articulate from the beginning of climate-policy efforts. The 1992 United Nations Framework Convention on Climate Change (UNFCCC) expressed the aim as preventing "dangerous anthropogenic interference in the climate system". Efforts to clarify the

meaning of 'dangerous' here have proved fruitless because science offers many different answers depending on which part of the climate system is under scrutiny, and each country has a different perspective<sup>2</sup>.

The 2009 and 2010 UNFCCC Conference of the Parties meetings, in Copenhagen and Cancun, Mexico, respectively, reframed the policy goal in more concrete terms: average global temperature. There was little scientific basis for the 2°C figure that was adopted, but it offered a simple focal point and was familiar from earlier discussions, including those by the IPCC, EU and Group of 8 (G8) industrial countries<sup>3</sup>. At the time, the 2°C goal sounded bold and perhaps feasible.

Since then, two nasty political problems have emerged. First, the goal is effectively unachievable<sup>4</sup>. Owing to continued failures to mitigate emissions globally, rising emissions are on track to blow through this limit eventually. To be sure, models show that it is just possible to make deep planet-wide cuts in emissions to meet the goal<sup>5</sup>. But those simulations make heroic assumptions — such as almost immediate global cooperation and widespread availability

ILLUSTRATION BY DAVID PARKINS



of technologies such as bioenergy carbon capture and storage methods that do not exist even in scale demonstration<sup>2</sup>.

Because it sounds firm and concerns future warming, the 2°C target has allowed politicians to pretend that they are organizing for action when, in fact, most have done little. Pretending that they are chasing this unattainable goal has also allowed governments to ignore the need for massive adaptation to climate change.

Second, the 2°C goal is impractical. It is related only probabilistically to emissions and policies, so it does not tell particular governments and people what to do. In other areas of international politics, goals have had a big effect when they have been translated into concrete, achievable actions<sup>6</sup>. For example, the eight Millennium Development Goals (MDGs) adopted by the United Nations in 2000 were effective when turned into 21 targets and 60 detailed indicators — measurable, practical and connected to what governments, non-governmental and aid organizations and others could do<sup>7</sup>.

### TROUBLING PAUSE

The scientific basis for the 2°C goal is tenuous. The planet's average temperature has barely risen in the past 16 years (see 'Heat exchange'). But other measures show that radiative forcing — the amount by which accumulating greenhouse gases in the atmosphere are perturbing the planet's energy balance — is accelerating<sup>8</sup>.

The Arctic, for example, has been warming rapidly. High-latitude climates are more sensitive than the planet as a whole. Amplifications in the Arctic might be causing extreme weather in middle latitudes<sup>9</sup>.

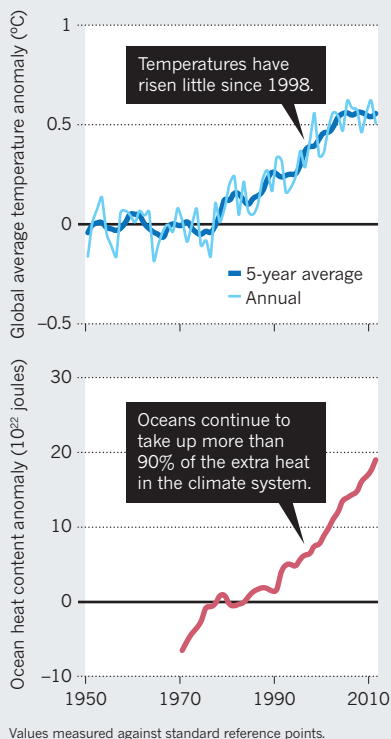
How could human stresses on the climate be rising faster even as global surface temperatures stay flat? The answer almost certainly lies in the oceans. The oceans are taking up 93% of the extra energy being added to the climate system, which is stoking sea-level rise and other climate impacts.

A single index of climate-change risk would be wonderful. Such a thing, however, cannot exist. Instead, a set of indicators is needed to gauge the varied stresses that humans are placing on the climate system and their possible impacts. Doctors call their basket of health indices vital signs. The same approach is needed for the climate.

The best indicator has been there all along: the concentrations of CO<sub>2</sub> and the other greenhouse gases (or the change in radiative forcing caused by those gases). Such parameters are already well measured through a network of international monitoring stations. A global goal for average concentrations in 2030 or 2050 must be agreed on and translated into specific emissions and policy efforts, updated periodically, so that individual governments can see clearly how

### HEAT EXCHANGE

Deep ocean waters have continued to become warmer despite global average temperature flattening off in the past 16 years.



their actions add up to global outcomes.

Some pollutants that perturb the climate, such as methane or soot, have huge regional and local variations, and important uncertainties remain about the link between human emissions and measured concentrations. Policy initiatives are gaining momentum that would improve measurement and control of those warming agents. For example, the Climate and Clean Air Coalition is a group of countries focused on ways to cut emissions of short-lived climate pollutants.

Policy-makers should also track ocean heat content and high-latitude temperature. Because energy stored in the deep oceans will be released over decades or centuries, ocean heat content is a good proxy for the long-term risk to future generations and planetary-scale ecology. High-latitude temperatures, because they are so sensitive to shifts in climate and they drive many tangible harms, are also useful to include in the planetary vital signs<sup>9</sup>.

### CHART A PATH

What is ultimately needed is a volatility index that measures the evolving risk from extreme events — so that global vital signs can be coupled to local information on what people care most about. A good start would be to track the total area during the year in which conditions stray by three standard deviations from the local and seasonal mean<sup>10</sup>.

The window of opportunity for improving

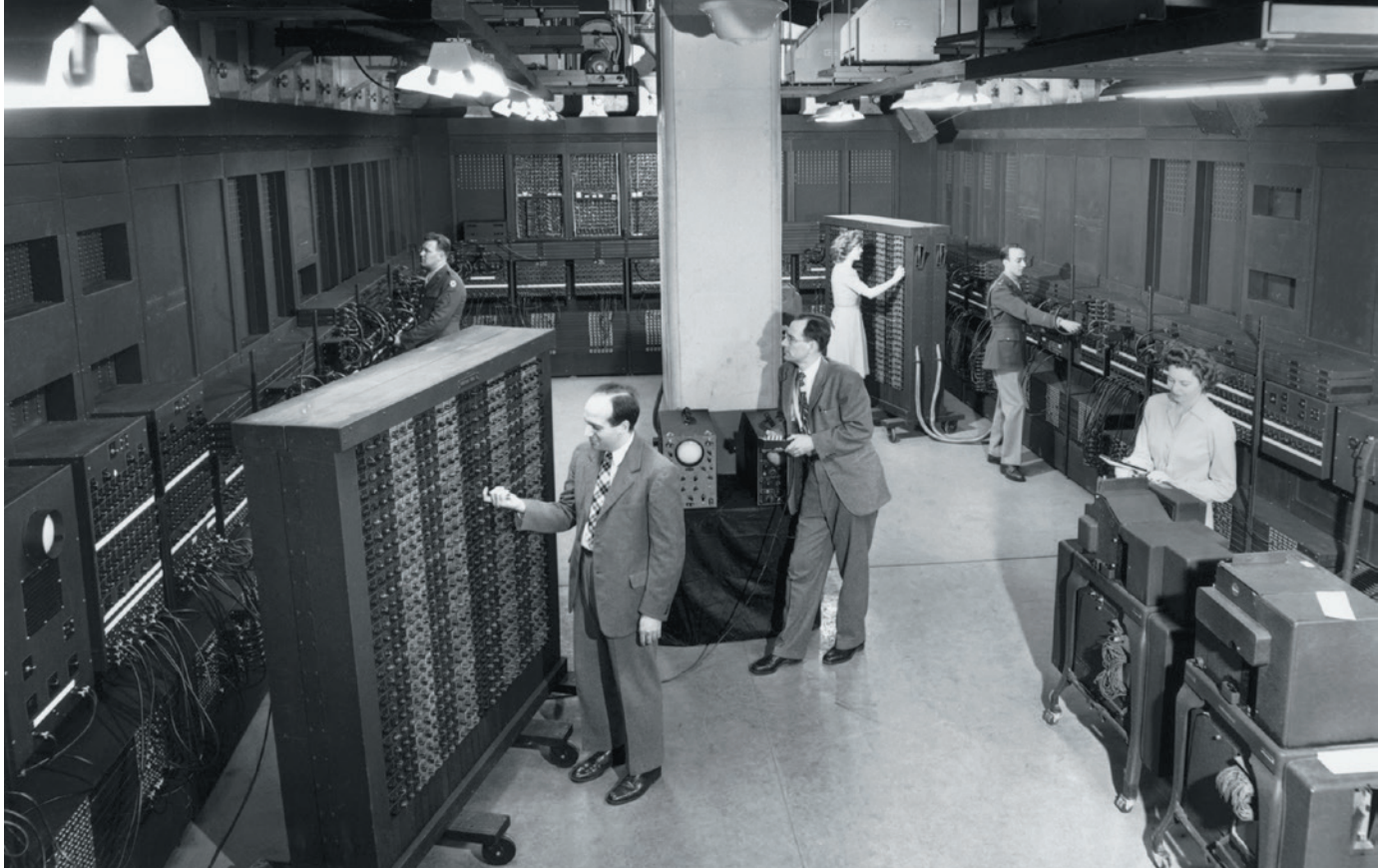
goal-setting is open. This autumn, a big push on climate policy begins — with the aim of crafting a new global agreement by late 2015 at the UNFCCC's Conference of the Parties in Paris. Getting serious about climate change requires wrangling about the cost of emissions goals, sharing the burdens and drawing up international funding mechanisms. But diplomats must move beyond the 2°C goal. Scientists must help them to understand why, and what should replace it.

New indicators will not be ready for the Paris meeting, but a path for designing them should be agreed there. Such a clear international mandate would spur research on indicators of planetary health, just as the United Nations' Millennium Summit on extreme poverty gave political momentum to the MDGs. The Paris agreement should call for an international technical conference on how to turn today's research measurements into tomorrow's planetary vital signs.

The public needs to understand what it is being asked to pay for. On this score, 'CO<sub>2</sub> concentration' or 'ocean heat content' are not nearly effective as 'temperature' in conveying to the person in the street what is at risk. Yet patients have come to understand that doctors must track many vital signs — blood pressure, heart rate and body mass index — to prevent illness and inform care. A similar strategy is now needed for the planet. ■

**David G. Victor** is professor of international relations at the School of International Relations and Pacific Studies, University of California, San Diego, La Jolla, California, USA. **Charles F. Kennel** is distinguished professor and director emeritus at the Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California, USA. e-mail: david.victor@ucsd.edu

1. Balmaseda, M. A., Trenberth, K. E. & Källén, E. *Geophys. Res. Lett.* **40**, 1754–1759 (2013).
2. Victor, D. G. et al. in *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Edenhofer, O. et al.) (Cambridge Univ. Press, 2014).
3. Randalls, S. *WIREs Clim. Change* **1**, 598–605 (2010).
4. Victor, D. G. *Global Warming Gridlock: Creating More Effective Strategies for Protecting the Planet* (Cambridge Univ. Press, 2011).
5. Clarke, L. et al. in *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Edenhofer, O. et al.) (Cambridge Univ. Press, 2014).
6. Levy, M. A. in *Institutions for the Earth: Sources of Effective International Environmental Protection* (eds Hass, P. M., Keohane, R. O. & Levy, M. A.) 75–132 (MIT Press, 1993).
7. Bourguignon, F. et al. in *Equity and Growth in a Globalizing World* 17–40 (eds Kanbur, R. & Spence, M.) (World Bank, 2010).
8. Hansen, J., Kharecha, P. & Sato, M. *Environ. Res. Lett.* **8**, 011006 (2013).
9. Cohen, J. et al. *Nature Geosci.* **7**, 627–637 (2014).
10. Hansen, J., Sato, M. & Ruedy, R. *Proc. Natl Acad. Sci. USA* **109**, E24150–E2423 (2012).



BETTMANN/CORBIS

The Electronic Numerical Integrator And Computer (ENIAC) and its co-designers, J. Presper Eckert (front left) and John Mauchly (at pillar).

## COMPUTING HISTORY

# Geeks, Inc.

Jennifer Light enjoys a chronicle of the collaborations that conjured the digital realm.

Pulitzer-prizewinning historian James McPherson has described history as a dialogue between past and present: everything from social movements to scientific discovery prompts historians to ask new questions about old events. Accordingly, our era's attentiveness to networks of people and technologies has triggered an outpouring of research that places innovation in collaborations, rather than crediting the lone genius who dominated older studies.

Walter Isaacson's *The Innovators* exemplifies this newer approach. Isaacson, whose career as best-selling author was built on the 'genius' biography — such as the 2011 *Steve Jobs* (Simon & Schuster) — now aims to present the definitive history of the digital revolution with teamwork at its core. After kicking off with the visionary work of nineteenth-century pioneers Ada Lovelace and Charles Babbage on the concept of a mechanical computer, Isaacson focuses squarely on the twentieth and twenty-first centuries. His book synthesizes and reworks academic studies in computing history, and draws on new interviews with technology pioneers such as Bill Gates — covering everything from his teenage adventures in the Lakeside Programming Group in Seattle, Washington,

### **The Innovators: How a Group of Inventors, Hackers, Geniuses, and Geeks Created the Digital Revolution**

WALTER ISAACSON

Simon & Schuster: 2014.

to the founding of Microsoft. This is a mostly sunny account, focused on the excitement of invention rather than the sometimes darker consequences of using digital technologies.

Each chapter is organized around a class of technology such as the computer, the transistor or the web. Isaacson uses the framework of collaboration to assess successes and failures. In 1939, for instance, US physicist John Atanasoff invented a machine that some deem to have been the first electronic digital computer. It was overshadowed by the mid-1940s debut of the US Army's Electronic Numerical Integrator And Computer (ENIAC) — credited by Isaacson with being the first general-purpose electronic computer. Atanasoff, he argues, failed to make his ideas pay owing to his isolation in Iowa and lack of access to the intellectual and financial resources of a team.

Atop the collaboration narrative, Isaacson layers another theory of how innovation happens, more in tune with his work on genius. Most of the gifted individuals in the creative teams he studies were, in his interpretation, as

much influenced by the arts or humanities as by technical training. For example, William Shockley, co-inventor of the transistor, "grew up with a love of both art and science". US computer scientist J. C. R. Licklider, in Isaacson's estimation "the single most important person in creating the Internet", was an art maven and collector.

*The Innovators* also adds a dash of anecdote. Who wouldn't be charmed to know that ENIAC project co-director J. Presper Eckert was, as a student, responsible for the 'Osculometer' — an electrical device to "measure the passion and romantic electricity of a kiss"? Isaacson also thoughtfully introduces basic themes in the history of invention. For example, he lays out how a tension between secrecy and openness characterizes much development. So on the one hand, there is the hacker ethos of the Homebrew Computer Club ("Software wants to be free"), and on the other, developers seeking compensation for their intellectual property. Isaacson also probes how technologies often evolve incrementally rather than arriving in a eureka moment, and how even in an era of electronic communication, places matter as much as environments for creative collisions of ideas. California's Silicon Valley, for instance, is now



a global model for ‘innovation districts.’ The book also provides simple explanations of basic technical concepts, such as analogue versus digital, and packet switching.

As a contribution to understanding of the digital revolution, however, *The Innovators* suffers from the same limitations that currently vex academic computing history: it is too white, male and insular. Aside from a nod to British computing pioneer Alan Turing, largely missing here are the stories of innovation beyond the United States — as are reflections on how non-digital games and alternative media might have shaped the design of digital technologies or the sense of what they might be used to achieve.

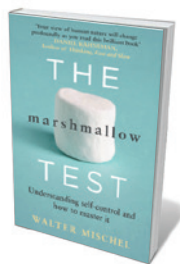
Questions about gender in particular demand further discussion. Isaacson begins and ends with Lovelace, whose “love of both poetry and math”, which primed her to see beauty in a computing machine, frames his assertions about the interconnectedness of art and science. Early chapters mention the contributions of female programming pioneers, including the six-woman ENIAC programming team. Yet after that, women mostly drop out of the action, and we encounter stories of how an early sales brochure for Atari games featured a woman in a sheer nightgown, hired “from the topless bar down the street”, or how Licklider routinely slipped photos of beautiful women into colleagues’ presentations.

Did female innovators remain in this hostile environment, forgotten to history? Did they find it so intolerable that they left? Both stories would be instructive today in light of widely recognized gender problems in US technology firms, as well as in Isaacson’s ambition to “explore the social and cultural forces that provide the atmosphere for innovation”.

The weakest aspect of the book is Isaacson’s attempt to link the arts to innovation, which he never quite backs up. Yet scholars have already shown the value in that view, and discussions about STEAM (science, technology, engineering, arts and mathematics), rather than STEM, are rife in educational institutions. The Massachusetts Institute of Technology in Cambridge — my employer — demands that science and engineering undergraduates take one-quarter of their courses in the arts, humanities or social sciences, on the basis of their recognized relevance to students’ work. Isaacson’s gift for digesting scholarly materials and making them accessible would have been well applied here, beyond the borders of computing history, to make the case for a multidisciplinary education. ■

**Jennifer Light** is professor of science, technology and society at the Massachusetts Institute of Technology in Cambridge, where she teaches courses on the histories of innovation and computing.  
e-mail: [jslight@mit.edu](mailto:jslight@mit.edu)

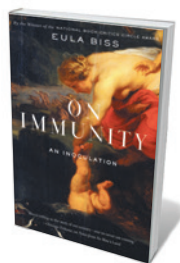
## Books in brief



### The Marshmallow Test: Understanding Self-control and How To Master It

Walter Mischel BANTAM (2014)

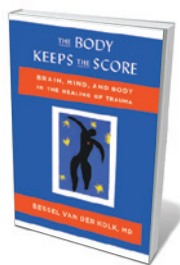
In our go-faster era, extreme impulsivity — from trolling to air rage — seems to be on the rise. So it is an apt moment for psychologist Walter Mischel to recap his much-cited “marshmallow test”, which examines children’s capacity for delaying gratification as an indicator of emotional balance in maturity. Mischel takes us beyond the experiment into deep research on “delay ability”, his formulation of “hot” and “cool” cognition, speculation on the role of genetics, and the implications of his work for public policy.



### On Immunity: An Inoculation

Eula Biss GRAYWOLF (2014)

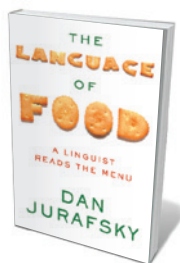
Our long and intimate coexistence with viruses is less battle than balancing act, avers essayist Eula Biss. In this quietly impassioned call for responsible childhood immunization, Biss explores the currents of humanity’s uneasy relationship with these microscopic hordes, interweaving science, myth and history with her own fraught parental experience. The word inoculate was originally used to describe plant grafting, she notes. Now, it signifies grafting disease “to the rootstock of the body”. As Biss reminds us, immunization must be effectively communal, “a garden that we tend together”.



### The Body Keeps the Score: Brain, Mind, and Body in the Healing of Trauma

Bessel van der Kolk VIKING ADULT (2014)

War zones may be nearer than you think, as the 25% of US citizens raised with alcoholic relatives might attest. Psychiatrist Bessel van der Kolk argues, moreover, that severe trauma is “encoded in the viscera” and demands tailored approaches that enable people to experience deep relief from rage and helplessness. In a narrative packed with decades of findings and case studies, he traces the evolution of treatments from the ‘chemical coshes’ of the 1970s to neurofeedback, mindfulness and other nuanced techniques.



### The Language of Food: A Linguist Reads the Menu

Dan Jurafsky W. W. NORTON (2014)

When Dan Jurafsky enters a restaurant, menu scribes beware: this linguist will pick at the wording even as he savours (or deplores) the dish. In his study probing how foods and their names co-evolved, Jurafsky crafts a gastronomic atlas. We discover how Peruvian ceviche and vinegary British fish and chips can be traced back to *sikbāj*, a sweet-and-sour stew from sixth-century Persia. We marvel at how a fermented-fish sauce from southern China is the progenitor of all-American ketchup. And we find an unexpected chemical connection between ice cream and fireworks. Deliciously erudite.



### The Imaginary App

Edited by Paul D. Miller and Svitlana Matviyenko MIT PRESS (2014)

Are mobile apps an “oscillator between the imaginary and the realised”, or “charming junkware”? Multimedia artist Paul D. Miller (also known as DJ Spooky, That Subliminal Kid) and media scholar Svitlana Matviyenko explore this vaporous realm with contributors including Björk collaborator Scott Snibbe. The theory-laced result is for the digital devotee, but the authors’ apps, real and speculative, can be great fun; the optical illusion in Anna Munster’s Transparent Screen app, for instance, allows you to “text and walk without fear”. **Barbara Kiser**



From albatrosses to us, the interaction of 'neighbourly' gene networks drives evolution.

# GENETICS

## The neighbourly nature of evolution

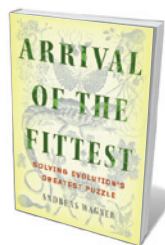
**Mark Pagel** relishes an analysis of how natural selection riffles through life's immense genetic library.

**Y**ou inhabit something of a miracle, in engineering terms. Your body consists of trillions of cells, woven together into something whose complexity far outstrips that of the most sophisticated objects our best engineers can produce, from computers and skyscrapers to space shuttles. A relatively simple outer form belies a teeming society of chemical reactions and protein engineering. This must maintain itself within strict temperature and physiological limits while enduring a complex and frequently unpredictable external environment. And, to achieve its long lifespan, it must avoid the sort of catastrophic breakdown that plagues human-engineered objects.

All the breathtaking innovation required to produce this complexity rests on two pillars of evolution that are, for the most part, either ignored or unappreciated. These are robustness and evolvability, which together

grant what evolutionary biologist Andreas Wagner calls "innovability" in his engaging and intelligent *Arrival of the Fittest*. Wagner's message is that these two foundation stones of evolution exist because of an unexpected and remarkable degree of neighbourliness (not his term) that seems to characterize life — a neighbourliness that allows species to innovate more rapidly and successfully than previously imagined.

Think of a rigid, riveted steel girder. It is, in many respects, a robust object, able to bear weight and resist high temperatures.



**Arrival of the Fittest: Solving Evolution's Greatest Puzzle**  
ANDREAS WAGNER  
Current/Oneworld:  
2014.

But it is not evolvable — there is nothing it can be but a girder. Now think of the most evanescent thing you can, perhaps a wisp of smoke in a breeze. It is highly evolvable — it can change — but it is not at all robust. The wonder of you and me and albatrosses is that we are not only robust, but also evolvable. Equally wonderful is that life as we know it would not be possible any other way.

Here is why. To get from simple replicating molecules through to single-celled organisms such as bacteria and eventually on to complex and ungainly multicellular organisms like giant squid, natural selection has had to search through a vast library of varieties and combinations of genes. Now, imagine you are in the squid section of the library and you want to make an albatross. Every step along the way has to be something that works: it has to be a competitive organism.

Wagner has discovered what makes this search possible. It is good neighbours, and lots of them. The genes that make our bodies typically do not act alone. Instead, they form large and complex networks that interact to produce metabolisms, tissues and organs. Wagner has built computer models of these networks in which he randomly alters some feature, mimicking *in silico* the sort of random mutation that natural selection relies on. He then asks whether the mutated network as a whole can still perform the job it was designed to do.

Overwhelmingly, the answer is yes, and it is this insensitivity to random change that makes biology robust to mutations and mishaps, and evolvable. Even better, Wagner finds that he does not have to travel very far along these mutational pathways before he encounters new neighbourhoods, where the networks produce different products. For instance, a network that can consume glucose might lie near one that can consume other fuels, such as acetate. Wagner thinks that these features of gene networks are repeated in proteins, metabolisms and the basic chemistry of cells. *In vivo* studies back him up.

This offers an answer to one of the most fundamental questions of evolution: how has natural selection had time to search the almost limitless library of life? The answer, posits Wagner, is that it does not usually have to search very far: squid and albatrosses are closer neighbours than we might have expected. *Arrival of the Fittest* will give you a new appreciation of the sheer improbability, but also the plausibility, of the diversity of life. ■

**Mark Pagel** is professor of evolutionary biology at the University of Reading, UK, and external professor at the Santa Fe Institute in New Mexico. He is the author of *Wired for Culture*.  
e-mail: m.pagel@reading.ac.uk



# Correspondence

## Research integrity guidelines in Japan

It is to be hoped that Japan's new guidelines for research integrity, released recently by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), will curb research misconduct (see T. Tanimoto *et al. Nature* **512**, 371; 2014).

Institutions in Japan have previously tended to avoid taking responsibility for misconduct by their scientists. Under the revised guidelines, a research institution must take appropriate measures against any scientist who is found guilty of data manipulation or fabrication, for example. Should it fail to do so, MEXT will cut its research budget.

MEXT has already reduced RIKEN's requested budget for next year by nearly 20% (¥12.1 billion; US\$111 million) as a penalty for inefficient handling of the two 'STAP' stem-cell papers published and subsequently retracted this year (see *Nature* **511**, 112; 2014). **Masanori Wada** Tokyo Institute of Technology, Japan.  
wada.m.ae@m.titech.ac.jp

## Lung-cancer screens now worth the cost

False-positive results from computed-tomography (CT) scans were a cause for concern in the 2011 US National Lung Screening Trial (see *Nature* **513**, S4–S6; 2014). But false-positives have now been cut significantly owing to improved imaging technology and more-refined screening protocols (see B. J. McKee *et al. J. Am. Coll. Radiol.* <http://dx.doi.org/10.1016/j.jacr.2014.08.002>; 2014).

Contrary to your implication, an actuarial analysis indicates that CT lung scanning is cost-effective (see B. Pyenson *et al. Am. Health Drug Benefits* **7**, 272–282; 2014) in the US population covered by the health-insurance programme

Medicare. Most members screened are aged over 65.

The study finds that the average monthly cost of CT lung scanning per Medicare member is just US\$1 (the equivalent screening cost for breast and colorectal cancers is \$2.50 and \$1.40, respectively). This latest cost-benefit analysis is consistent with other peer-reviewed research proving that lung-cancer screening is cost-effective for Medicare and for private payers too.

As James Mulshine, a translational-medicine specialist at Rush University in Chicago, Illinois, pointed out in a 2010 Lung Cancer Alliance statement, through screening, "we have the opportunity to realize the greatest single reduction of cancer mortality in the history of the war on cancer" (see [go.nature.com/vs2smt](http://go.nature.com/vs2smt)).

**Laurie Fenton-Ambrose** Lung Cancer Alliance, Washington DC, USA.

rryan@powelltate.com

**Ella A. Kazerooni** American College of Radiology, Washington DC, USA.

## Protect privacy of mobile data

The use of new data sources to model humans' behavioural responses to climate change (see P. Palmer and M. Smith *Nature* **512**, 365–366; 2014) raises methodological and ethical issues.

The authors do not mention the importance of call-detail records (CDRs), normally collected by mobile-phone operators for commercial purposes. Compared with data collected by smartphones from global-positioning systems, CDRs have lower location accuracy and differ in their potential for modelling and privacy risks.

The poorest communities in low- and middle-income settings should be at the centre of modelling efforts because

they are among the most vulnerable to climate change.

These populations are unlikely to use satellite navigation and social media, and they might not conform to human-mobility models derived from commuting patterns in wealthy countries. In these regions, CDRs can provide important insight — if appropriate privacy protections are in place (see, for example, A. Wesolowski *et al. Science* **338**, 267–270; 2012).

The rights of the individual to control their private data and the needs of researchers and policy-makers to access data for societal good create tensions that are central to the effective modelling of human behaviour (see, for example, Y. de Montjoye *et al. Sci. Rep.* **3**, 1376; 2013). These call for new regulatory and institutional review board processes.

**Caroline O. Buckee\*** Harvard School of Public Health, Boston, Massachusetts, USA; and Flowminder Foundation.

[cbuckee@hsph.harvard.edu](mailto:cbuckee@hsph.harvard.edu)  
\*On behalf of 5 correspondents (see [go.nature.com/pezxek](http://go.nature.com/pezxek) for full list.)

## Invest in renewable energy in Tibet

Tibet's fragile environment is being damaged by a paucity of energy, as well as by pollutants and litter (see *Nature* **512**, 240–241; 2014). Greater investment could unleash the region's huge potential to produce renewable energy.

Access to fossil fuels is extremely limited in Tibet, particularly in rural areas. Biomass — including manure, firewood and crop residues — is largely used instead, making up two-thirds of total energy use. However, this degrades forest and grassland and causes indoor pollution (see G. Liu *et al. Renew Sust. Energ. Rev.* **12**, 1890–1908; 2008).

Tibet has abundant resources

for renewable energy (including solar, wind and geothermal) owing to its complex topography and widely varying climate (see L. Shen *et al. Environ. Manage.* **46**, 539–554; 2010).

These resources remain mostly untapped, however, because of the high cost of exploitation, unevenly distributed settlements, lack of local infrastructure, and inadequate maintenance and knowledge.

**Gang Liu** Norwegian University of Science and Technology, Trondheim, Norway.

[geoliugang@gmail.com](mailto:geoliugang@gmail.com)

**Mario Lucas** Bochum University of Applied Sciences, Bochum, Germany.

## Focus on positive features of ageing

Ageing is not just a linear physiological decline (see L. Fontana *et al. Nature* **511**, 405–407; 2014). Research into its more positive features could lead to a better-quality and longer life.

Personal resources such as optimism, resilience and engagement are integral to ageing well (see, for example, T. D. Cosco *et al. BMJ Open* **3**, e002710; 2013).

Investigating the contribution of psychosocial strengths to positive ageing in human models would provide further insight into, and complement, physiological animal models.

**Theodore D. Cosco**, **Carol Brayne** University of Cambridge, UK.

[tdc33@medschl.cam.ac.uk](mailto:tdc33@medschl.cam.ac.uk)

**Blossom C. M. Stephan** Newcastle University, UK.

### CONTRIBUTIONS

Correspondence may be submitted to [correspondence@nature.com](mailto:correspondence@nature.com) after consulting the author guidelines at [go.nature.com/cmchno](http://go.nature.com/cmchno).

Inhibition of demethylases by GSK-J1/J4

ARISING FROM L. Kruidenier *et al. Nature* **488**, 404–408 (2012); doi:10.1038/nature11262

The recent publication<sup>1</sup> of the first highly potent and specific inhibitor GSK-J1/J4 of the H3K27me3/me2-demethylases JMJD3/KDM6B and UTX/KDM6A provides a potential tool compound for this histone demethylase subfamily<sup>1</sup>. This inhibitor was used in tissue culture assays to conclude that the catalytic activities of the KDM6 proteins are required in inflammatory responses<sup>1</sup>; the generation of the inhibitor is intriguing, because it provides a strategy for generating sub-type-specific inhibitors of the 27-member jumonji family and for the future treatment of various types of disease<sup>2–6</sup>. Here we show that the inhibitor is not specific for the H3K27me3/me2-demethylase subfamily *in vitro* and in tissue culture assays. Thus, the inhibitor cannot be used alone for drawing conclusions regarding the specific role of H3K27me3/me2-demethylase

activity in biological processes or disease. There is a Reply to this Brief Communications Arising by Kruidenier, L. *et al. Nature* **514**, http://dx.doi.org/10.1038/nature13689 (2014).

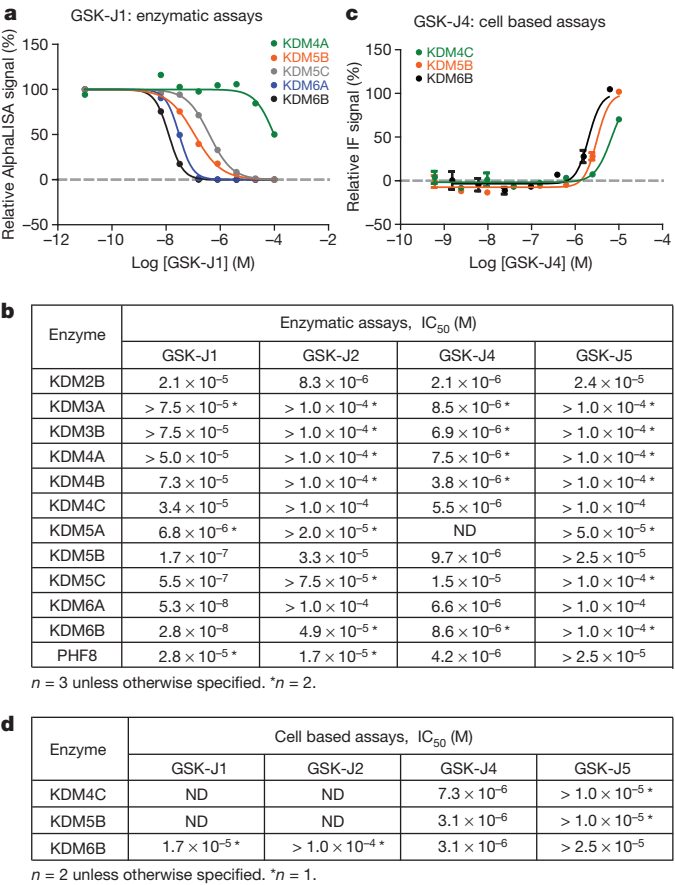
The jumonji demethylases are dependent on two cofactors, 2-oxoglutarate (also known as  $\alpha$ -ketoglutarate) and  $\text{Fe}^{2+}$  for enzymatic activity. The compound published by Kruidenier *et al.*<sup>1</sup>, GSK-J1 is a competitive inhibitor of the two cofactors, but not of the substrate, with a half-maximum inhibitory concentration ( $\text{IC}_{50}$ ) of 60 nM towards KDM6B as measured in an AlphaScreen assay. By performing *in vitro* assays on a number of other jumonji demethylases, including the closely related JMJD2/KDM4 subfamily and 160 other proteins, Kruidenier *et al.*<sup>1</sup> concluded that GSK-J1 is specific for the KDM6 subfamily. However, we noted that GSK-J1 was not tested on the JARID1/KDM5 subfamily, which contains the four demethylases with the closest homology in the catalytic domain to KDM6B and KDM6A (ref. 3). As shown in Fig. 1a, b, we tested the inhibitory activity of GSK-J1 towards 12 different jumonji demethylases. In agreement with the published data<sup>1</sup>, our results show that GSK-J1 is a highly potent inhibitor of KDM6B and KDM6A. Moreover, and also in agreement with Kruidenier *et al.*<sup>1</sup>, the other tested demethylases, except for KDM5B and KDM5C, were only marginally or not significantly inhibited *in vitro*. However, our results show that GSK-J1 only is fivefold to tenfold more potent towards KDM6B and KDM6A as compared to KDM5B and KDM5C. As a control for these experiments, we used GSK-J2, an isomer of GSK-J1 that does not have any specific activity<sup>1</sup>. Taken together, these results show that GSK-J1 is a potent inhibitor of jumonji proteins with activity towards H3K27me3/me2 (KDM6) and H3K4me3/me2 (KDM5) *in vitro*.

The highly polar GSK-J1 compound is restricted from entering into cells, and Kruidenier *et al.*<sup>1</sup> therefore changed the acid group in GSK-J1 and GSK-J2 to an ester, thereby generating GSK-J4 and GSK-J5, respectively<sup>1</sup>. In a mass-spectrometry based *in vitro* assay, GSK-J4 was shown to have an  $\text{IC}_{50} > 50 \mu\text{M}$ <sup>1</sup>. In a more sensitive AlphaLISA assay, we found that GSK-J4 has half-maximum inhibitory concentration ( $\text{IC}_{50}$ ) towards KDM6B and KDM6A of 8.6  $\mu\text{M}$  and 6.6  $\mu\text{M}$ , respectively (Fig. 1b). GSK-J4 was also found to inhibit the catalytic activity of the other tested demethylases with similar potency (Fig. 1b). Kruidenier *et al.*<sup>1</sup> did not report on the  $\text{IC}_{50}$  value of GSK-J4 towards different jumonji demethylases in transfected cells, however, they showed an  $\text{IC}_{50}$  value of 9  $\mu\text{M}$  towards the production of  $\text{TNF-}\alpha$  in lipopolysaccharide-stimulated macrophages. We tested the inhibitory effect of the four GSK compounds in cells transfected with KDM6B, KDM5B and KDM4C, respectively, and as shown in Fig. 1c, d, GSK-J4 shows very similar  $\text{IC}_{50}$  values towards the 3 demethylases, representing 3 different subfamilies. Taken together, our results show that GSK-J1 and GSK-J4 inhibit demethylases in addition to KDM6B and KDM6A. Therefore, this compound cannot be used alone for demonstrating a role for H3K27 demethylation in biological processes.

Methods

AlphaLISA assays were essentially performed as described in the protocol provided by the manufacturer (PerkinElmer). The enzymes used were: KDM2B (amino acids 1–650), KDM3A (amino acids 2–1,322), KDM3B (amino acids 842–1,761), KDM4A (amino acids 1–350), KDM4B (amino acids 2–500), KDM4C (amino acids 1–349), KDM5A (amino acids 1–1,090), KDM5B (amino acids 1–809), KDM5C (amino acids 2–1,560), KDM6A (amino acids 919–1,401), KDM6B (amino acids 1,043–1,643), and PHF8 (amino acids 1–1,024). Substrates and assay conditions can be provided upon request.

To measure the inhibitory activity of the tested compounds in cell-based assays, U2OS cells were transfected with epitope tagged versions of KDM6B (amino acids



**Figure 1 | GSK-J1/J4 inhibition of several histone demethylase subfamilies.** **a**, Assessment of the inhibitory potential of GSK-J1 towards the indicated jumonji enzymes by AlphaLISA based assays. **b**, Inhibitory potential of GSK-J1, GSK-J2, GSK-J4 and GSK-J5 towards the indicated enzymes as assessed by AlphaLISA based assays.  $\text{IC}_{50}$  values are indicated as means. The deviation of the mean was always less than twofold. ND, not determined. *n* equals the number of replicates. **c**, Assessment of inhibitory potential of GSK-J4 in cell-based assays in which the indicated enzymes were transfected, and their activity measured by induced loss of H3K27me2 (KDM6B), H3K4me2 (KDM5B) and H3K9me3 (KDM4C). **d**, Inhibitory activity of the indicated compounds towards the indicated enzymes in cell-based assays.  $\text{IC}_{50}$  values are indicated as mean, and the deviation from the mean was always less than twofold. ND, not determined.



1,026–1,682), KDM5B (amino acids 1–752) or full length KDM4C. Transfected cells were incubated with the indicated concentration of compounds, and the activity of the demethylase towards substrate in transfected cells was measured using antibodies specific for H3K27me2 (Abcam Ab24684), H3K4me2 (Milipore 07-030) and H3K9me3 (Abcam Ab8898).

**Bo Heinemann<sup>1</sup>, Jesper Morten Nielsen<sup>1</sup>, Heidi Rye Hudlebusch<sup>1</sup>, Michael J. Lees<sup>2</sup>, Dorthe Vang Larsen<sup>1</sup>, Thomas Boesen<sup>1</sup>, Marc Labelle<sup>1</sup>, Lars-Ole Gerlach<sup>1</sup>, Peter Birk<sup>1</sup> & Kristian Helin<sup>2,3,4</sup>**

<sup>1</sup>EpiTherapeutics Aps, Ole Maaløes Vej 3, 2200 Copenhagen, Denmark.

<sup>2</sup>Biotech Research and Innovation Centre (BRIC), University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen, Denmark.

<sup>3</sup>Centre for Epigenetics, University of Copenhagen, Ole Maaløes Vej 5, 2200 Copenhagen, Denmark.

<sup>4</sup>The Danish Stem Cell Center (DanStem), University of Copenhagen, Blegdamsvej 3, 2200 Copenhagen, Denmark.

email: kristian.helin@bric.ku.dk

## Kruidenier *et al.* reply

REPLYING TO B. Heinemann *et al.* *Nature* **514**, <http://dx.doi.org/10.1038/nature13688> (2014)

We welcome the accompanying Comment<sup>1</sup> by Heinemann *et al.*, in which the authors use an extensive panel of sensitive KDM assays to independently confirm our results<sup>2</sup> that GSK-J1 is a potent KDM6 inhibitor. Additionally, Heinemann *et al.*<sup>1</sup> demonstrate that GSK-J1 has some, albeit weaker, activity towards KDM5B and KDM5C, for which we only had preliminary data available at the time of our original publication. As our jumonji assay portfolio expands, we have continued to update the GSK-J1 activity profile on the SGC website (<http://www.thesgc.org/chemical-probes/GSKJ1>); this includes KDM5 inhibition activity by GSK-J1 similar to that reported by Heinemann. In conclusion, GSK-J1 remains the most selective KDM inhibitor yet disclosed and thus a valuable chemical tool.

Heinemann *et al.*<sup>1</sup> also show a broader, weak micromolar KDM inhibitory activity of the ester pro-drug version of GSK-J1, GSK-J4. GSK-J4 is not itself a chemical tool for direct KDM inhibition, but was designed specifically to enable efficient intracellular delivery of GSK-J1 into macrophages. In our work, the intracellular conversion of ester pro-drug is complete within 15 min after which levels of intracellular GSK-J4 are negligible ([GSK-J4] = 150 nM; [GSK-J1] = 11.8 µM). This renders the activity profile of GSK-J4 irrelevant and the biological effects in macrophages will be exclusively driven by the activity of GSK-J1. For other cell systems, it is essential to assess the ability to convert GSK-J4 to GSK-J1 before conducting and interpreting biological studies.

Despite the refinement of the selectivity profile of GSK-J1, our conclusion that KDM6 enzymatic activity is a key determinant of lipopolysaccharide responses in macrophages stands and was independently verified using short interfering RNA (siRNA) mediated knockdown of KDM6 enzymes. GSK-J1 remains a useful chemical probe for studying the catalytic function of KDM6 and the additional KDM5 activity may provide new opportunities for its use.

Received 23 January; accepted 25 June 2014.

1. Kruidenier, L. *et al.* A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. *Nature* **488**, 404–408 (2012).
2. Helin, K. & Dhanak, D. Chromatin proteins and modifications as drug targets. *Nature* **502**, 480–488 (2013).
3. Kooistra, S. M. & Helin, K. Molecular mechanisms and potential functions of histone demethylases. *Nature Rev. Mol. Cell Biol.* **13**, 297–311 (2012).
4. Arrowsmith, C. H., Bountra, C., Fish, P. V., Lee, K. & Schapira, M. Epigenetic protein families: a new frontier for drug discovery. *Nature Rev. Drug Discov.* **11**, 384–400 (2012).
5. Greer, E. L. & Shi, Y. Histone methylation: a dynamic mark in health, disease and inheritance. *Nature Rev. Genet.* **13**, 343–357 (2012).
6. Højfeldt, J. W., Agger, K. & Helin, K. Histone lysine demethylases as targets for anticancer therapy. *Nature Rev. Drug Discov.* **12**, 917–930 (2013).

**Author Contributions** B.H., J.M.N. and H.R.H. contributed equally to this manuscript. B.H., J.M.N., H.R.H., M.J.L. and D.V.L. performed experiments and analysed data.

T.B., M.L., L.-O.G., P.B. and K.H. analysed data. K.H. wrote the manuscript with input from the other authors.

**Competing Financial Interests** B.H., J.M.N., H.R.H., D.V.L., T.B., M.L., L.-O.G. and P.B. are employees of EpiTherapeutics Aps. K.H. is a co-founder, stockowner and consultant of EpiTherapeutics Aps. M.J.L. has no competing interests to declare.

doi:10.1038/nature13688

**Laurens Kruidenier<sup>1</sup>, Chun-wa Chung<sup>2</sup>, Zhongjun Cheng<sup>3</sup>, John Liddle<sup>1</sup>, KaHing Che<sup>4,5</sup>, Gerard Joberty<sup>6</sup>, Marcus Bantscheff<sup>6</sup>, Chas Bountra<sup>4</sup>, Angela Bridges<sup>2</sup>, Hawa Diallo<sup>1</sup>, Dirk Eberhard<sup>6</sup>, Sue Hutchinson<sup>2</sup>, Emma Jones<sup>2</sup>, Roy Katso<sup>2</sup>, Melanie Leveridge<sup>2</sup>, Palwinder K. Mander<sup>1</sup>, Julie Mosley<sup>2</sup>, Cesar Ramirez-Molina<sup>1</sup>, Paul Rowland<sup>2</sup>, Christopher J. Schofield<sup>4</sup>, Robert J. Sheppard<sup>1</sup>, Julia E. Smith<sup>1</sup>, Catherine Swales<sup>5</sup>, Robert Tanner<sup>2</sup>, Pamela Thomas<sup>2</sup>, Anthony Tumber<sup>4</sup>, Gerard Drewes<sup>6</sup>, Udo Oppermann<sup>4,5</sup>, Dinshaw J. Patel<sup>3</sup>, Kevin Lee<sup>1†</sup> & David M. Wilson<sup>1</sup>**

<sup>1</sup>Epinova DPU, Immuno-Inflammation Therapy Area, GlaxoSmithKline R&D, Medicines Research Centre, Gunnels Wood Road, Stevenage SG1 2NY, UK.

<sup>2</sup>Platform Technology and Science, GlaxoSmithKline R&D, Medicines Research Centre, Gunnels Wood Road, Stevenage SG1 2NY, UK.

email: chun-wa.h.chung@gsk.com

<sup>3</sup>Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10065, USA.

<sup>4</sup>Structural Genomics Consortium, University of Oxford, Old Road Campus, Roosevelt Drive, Headington OX3 7DQ, UK.

<sup>5</sup>Botnar Research Centre, NIHR Biomedical Research Unit, University of Oxford OX3 7LD, UK.

<sup>6</sup>Cellzome AG, Meyerhofstrasse 1, 69117 Heidelberg, Germany.

†Present address: Pfizer, Biotherapeutics R&D, 200 Cambridgepark Drive, Cambridge, Massachusetts 02140, USA.

1. Heinemann, B. *et al.* Inhibition of demethylases by GSK-J1/J4. *Nature* **514**, <http://dx.doi.org/10.1038/nature13688> (2014).
2. Kruidenier, L. *et al.* A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. *Nature* **488**, 404–408 (2012).

doi:10.1038/nature13689

# Another explanation for apparent epistasis

ARISING FROM G. Hemani *et al.* *Nature* **508**, 249–253 (2014); doi:10.1038/nature13005

Epistasis occurs when the effect of a genetic variant on a trait is dependent on genotypes of other variants elsewhere in the genome. Hemani *et al.* recently reported the detection and replication of many instances of epistasis between pairs of variants influencing gene expression levels in humans<sup>1</sup>. Using whole-genome sequencing data from 450 individuals we strongly replicated many of the reported interactions but, in each case, a single third variant captured by our sequencing data could explain all of the apparent epistasis. Our results provide an alternative explanation for the apparent epistasis observed for gene expression in humans. There is a Reply to this Brief Communication Arising by Hemani, G. *et al.* *Nature* **514**, <http://dx.doi.org/10.1038/nature13692> (2014).

Hemani *et al.*<sup>1</sup> identified 30 pairs of single nucleotide polymorphisms (SNPs; Table 1 in Hemani *et al.*<sup>1</sup>) that interacted to influence the expression of 19 different gene transcripts. These interactions were robust to adjustment for multiple testing and were replicated across two independent studies. Most of the replicated apparently interacting SNP pairs were associated with gene expression in *cis* and were located close to each other on the same chromosome (all <520 kilobases). We have previously shown that low levels of correlation due to linkage disequilibrium (LD) between variants can cause apparent allelic heterogeneity at an associated locus<sup>2</sup>. We therefore hypothesized that low levels of LD could explain the epistasis observed by Hemani *et al.*<sup>1</sup>.

To address this hypothesis, we used a combination of whole-genome sequence data and whole-blood gene expression traits in 450 individuals from the InCHIANTI study<sup>2</sup>. Gene expression levels were measured using a similar Illumina array (Human HT-12 v3.0) as Hemani *et al.*<sup>1</sup>

used for all of their discovery and replication analyses and we used the same analysis software (epiGPU<sup>3</sup>).

We first replicated the apparent interactions detected and replicated by Hemani *et al.* (11 of 17 *cis-cis* pairs and 3 of 11 *cis-trans* pairs with  $P < 0.05$ ; Table 1). Our lower success rate of replicating the *cis-trans* effects is consistent with their reported smaller effect sizes. We could not analyse two of the gene expression traits because either the probe or one of the SNPs failed quality control in our study. We next identified the single most strongly associated variant for each of the 17 gene expression traits from our whole-genome sequencing analysis. For 27 out of 28 SNP pairs the individual variant most strongly associated with gene expression in our data was more strongly associated than the 8 degrees of freedom (8 d.f.) full model formed from the pair of SNPs reported in Hemani *et al.* (Table 1). For all 17 putatively interacting pairs where both SNPs occurred on the same chromosome our more strongly associated variant was moderately correlated with both of the interacting SNPs (Table 2). These correlations occurred despite very low levels of LD between the two SNPs described by Hemani *et al.*

We next re-evaluated the evidence for interaction but this time corrected for the presence of our most strongly associated variant. The inclusion of our third variant removed any evidence for interaction (Table 1). This included the removal of apparently strong interactions involving *cis* variants for *MBNL1* and *TMEM149* (also known as *IGFLR1*), the two transcripts that account for all of the *cis-trans* interactions. Additionally, the most strongly associated variant for *MBNL1* occurs in the probe sequence used to detect expression of the gene, raising the possibility

**Table 1 | Results from running pairwise SNP interaction analyses on SNP pairs identified and replicated by Hemani *et al.*<sup>1</sup> and the results observed after conditioning on the most strongly associated additive *cis* variant identified in the InCHIANTI sequencing study (IncSeq)**

Cis/trans	Gene (chr)	SNP pairs from Hemani <i>et al.</i> Table 1			Two SNPs from Hemani <i>et al.</i>		Adjusted for IncSeq variant	
		SNP1 (chr)	SNP2 (chr)	IncSeq variant*	8 d.f. full model <i>P</i>	Interaction <i>P</i>	8 d.f. full model <i>P</i>	Interaction <i>P</i>
Cis	ADK (10)	rs2395095 (10)	rs10824092 (10)	10:75928933	$3.2 \times 10^{-19}$	$9.1 \times 10^{-04}$	0.99	0.86
Cis	ATP13A1 (19)	rs4284750 (19)	rs873870 (19)	19:19756073	$2.1 \times 10^{-05}$	$7.9 \times 10^{-03}$	0.87	0.64
Cis	C21ORF57 (21)	rs9978658 (21)	rs11701361 (21)	21:47703649	$3.8 \times 10^{-05}$	$7.2 \times 10^{-03}$	0.02	0.43
Cis	CSTB (21)	rs9979356 (21)	rs3761385 (21)	21:45201832	$6.2 \times 10^{-07}$	$8.3 \times 10^{-07}$	0.98	0.99
Cis	CTSC (11)	rs7930237 (11)	rs556895 (11)	11:88015717	$3.5 \times 10^{-15}$	$5.0 \times 10^{-06}$	$7.0 \times 10^{-08}$	0.04
Cis	FN3KRP (17)	rs898095 (17)	rs9892064 (17)	17:80678628	$2.8 \times 10^{-11}$	$2.9 \times 10^{-12}$	0.07	0.43
Cis	GAA (17)	rs11150847 (17)	rs12602462 (17)	17:78096086	0.09	0.15	0.22	0.34
Cis	HNRPH1 (5)	rs6894268 (5)	rs4700810 (5)	5:178978883	0.08	0.53	0.36	0.45
Cis	LAX1 (1)	rs1891432 (1)	rs10900520 (1)	1:203747772	$8.3 \times 10^{-08}$	$1.6 \times 10^{-04}$	0.27	0.52
Cis	MBNL1 (3)	rs16864367 (3)	rs13079208 (3)	3:152182577	$1.1 \times 10^{-07}$	$2.7 \times 10^{-06}$	0.41	0.16
Trans	MBNL1 (3)	rs7710738 (5)	rs13069559 (3)	3:152182577	$3.1 \times 10^{-05}$	$2.3 \times 10^{-02}$	0.05	0.02
Trans	MBNL1 (3)	rs2030926 (6)	rs13069559 (3)	3:152182577	$2.2 \times 10^{-05}$	$3.2 \times 10^{-02}$	0.19	0.21
Trans	MBNL1 (3)	rs2614467 (14)	rs13069559 (3)	3:152182577	$3.7 \times 10^{-04}$	0.24	0.47	0.55
Trans	MBNL1 (3)	rs218671 (17)	rs13069559 (3)	3:152182577	$1.4 \times 10^{-03}$	0.90	0.38	0.79
Trans	MBNL1 (3)	rs11981513 (7)	rs13069559 (3)	3:152182577	$1.6 \times 10^{-05}$	$1.6 \times 10^{-02}$	0.11	0.10
Cis	MBP (18)	rs8092433 (18)	rs4890876 (18)	18:74723459	$1.2 \times 10^{-02}$	0.05	0.67	0.28
Cis	NAPRT1 (8)	rs2123758 (8)	rs3889129 (8)	8:144684215	$6.8 \times 10^{-34}$	$6.2 \times 10^{-06}$	0.40	0.84
Cis	NCL (2)	rs7563453 (2)	rs4973397 (2)	2:232320581	0.09	0.10	0.85	0.71
Cis	PRMT2 (21)	rs2839372 (21)	rs11701058 (21)	21:47887791	$2.6 \times 10^{-15}$	$2.6 \times 10^{-04}$	0.52	0.30
Cis	SNORD14A (11)	rs2634462 (11)	rs6486334 (11)	11:17230389	$1.7 \times 10^{-05}$	0.37	0.41	0.17
Cis	TMEM149 (19)	rs807491 (19)	rs7254601 (19)	19:36234489	$3.0 \times 10^{-31}$	$2.9 \times 10^{-06}$	0.46	0.41
Trans	TMEM149 (19)	rs8106959 (19)	rs6926382 (6)	19:36234489	$3.2 \times 10^{-43}$	0.23	0.17	0.53
Trans	TMEM149 (19)	rs8106959 (19)	rs914940 (1)	19:36234489	$3.7 \times 10^{-42}$	0.62	0.39	0.71
Trans	TMEM149 (19)	rs8106959 (19)	rs2351458 (4)	19:36234489	$3.5 \times 10^{-42}$	0.30	0.53	0.46
Trans	TMEM149 (19)	rs8106959 (19)	rs6718480 (2)	19:36234489	$6.1 \times 10^{-42}$	0.44	0.57	0.69
Trans	TMEM149 (19)	rs8106959 (19)	rs1843357 (8)	19:36234489	$4.0 \times 10^{-41}$	0.44	0.91	0.73
Trans	TMEM149 (19)	rs8106959 (19)	rs9509428 (13)	19:36234489	$3.3 \times 10^{-42}$	0.09	0.69	0.39
Cis	VASP (19)	rs1264226 (19)	rs2276470 (19)	19:46033382	0.12	0.81	0.71	0.56

Data was available for 28 of the 30 interactions reported by Hemani *et al.*<sup>1</sup>. Both the full model and interaction associations for the Hemani *et al.* SNPs are completely removed on adjustment for the additive effect of our single most associated variant.

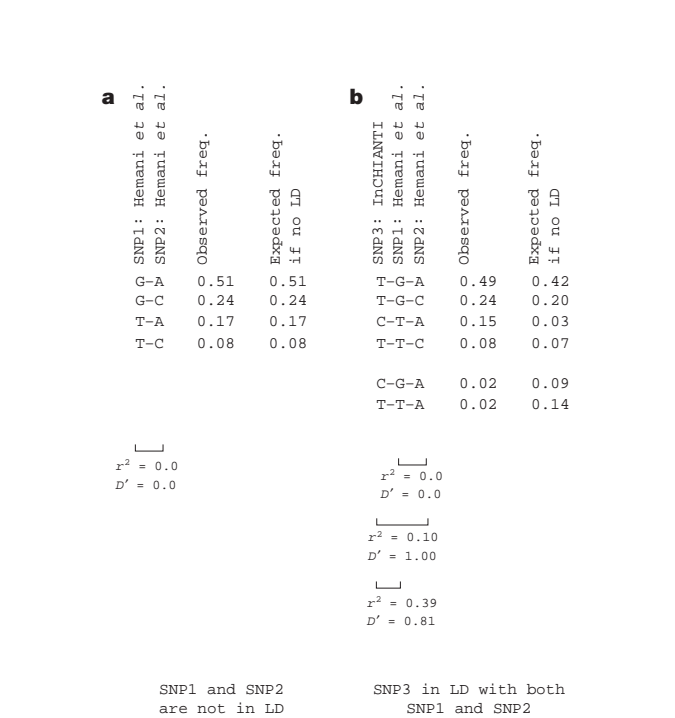
\* IncSeq variant is the most strongly associated additive variant with probe levels in *cis* ( $\pm 1$  Mb probe start site).



**Table 2 | Linkage disequilibrium measures between SNP pairs identified by Hemani *et al.*<sup>1</sup> and the most strongly associated *cis* variant identified in the InCHIANTI sequencing study**

<i>Cis/trans</i>	Gene (chr)	SNP pairs from Hemani <i>et al.</i> Table 1		IncSeq variant*	Linkage disequilibrium between variants		
		SNP1 (chr)	SNP2 (chr)		SNP1 – SNP2 $r^2/D'$	SNP1 – IncSeq $r^2/D'$	SNP2 – IncSeq $r^2/D'$
<i>Cis</i>	ADK (10)	rs2395095 (10)	rs10824092 (10)	10:75928933	0/0.01	0.39/0.81	0.1/1
<i>Cis</i>	ATP13A1 (19)	rs4284750 (19)	rs873870 (19)	19:19756073	0.01/0.11	0.07/0.9	0.04/0.82
<i>Cis</i>	C21ORF57 (21)	rs9978658 (21)	rs11701361 (21)	21:47703649	0.02/0.19	0.02/0.2	0.02/0.21
<i>Cis</i>	CSTB (21)	rs9979356 (21)	rs3761385 (21)	21:45201832	0.04/0.23	0.05/0.25	0.14/0.38
<i>Cis</i>	CTSC (11)	rs7930237 (11)	rs556895 (11)	11:88015717	0/0.07	0.22/0.9	0.11/0.94
<i>Cis</i>	FN3KRP (17)	rs898095 (17)	rs9892064 (17)	17:80678628	0/0.04	0.01/0.12	0.05/0.27
<i>Cis</i>	GAA (17)	rs11150847 (17)	rs12602462 (17)	17:78096086	0.01/0	0.3/1	0.11/0.94
<i>Cis</i>	HNRPH1 (5)	rs6894268 (5)	rs4700810 (5)	5:178978883	0.02/0.23	0.05/0.42	0.3/0.63
<i>Cis</i>	LAX1 (1)	rs1891432 (1)	rs10900520 (1)	1:203747772	0.03/0.23	0.21/0.51	0.05/0.29
<i>Cis</i>	MBNL1 (3)	rs16864367 (3)	rs13079208 (3)	3:152182577	0.08/0.42	0.13/0.62	0.06/1
<i>Trans</i>	MBNL1 (3)	rs7710738 (5)	rs13069559 (3)	3:152182577	NA	NA	0.44/1
<i>Trans</i>	MBNL1 (3)	rs2030926 (6)	rs13069559 (3)	3:152182577	NA	NA	0.44/1
<i>Trans</i>	MBNL1 (3)	rs2614467 (14)	rs13069559 (3)	3:152182577	NA	NA	0.44/1
<i>Trans</i>	MBNL1 (3)	rs218671 (17)	rs13069559 (3)	3:152182577	NA	NA	0.44/1
<i>Trans</i>	MBNL1 (3)	rs11981513 (7)	rs13069559 (3)	3:152182577	NA	NA	0.44/1
<i>Cis</i>	MBP (18)	rs8092433 (18)	rs4890876 (18)	18:74723459	0.04/0.22	0.11/0.43	0.21/0.62
<i>Cis</i>	NAPRT1 (8)	rs2123758 (8)	rs3889129 (8)	8:144684215	0.03/0.17	0.4/0.96	0.06/0.68
<i>Cis</i>	NCL (2)	rs7563453 (2)	rs4973397 (2)	2:232320581	0.04/0.25	0.29/0.83	0.16/0.76
<i>Cis</i>	PRMT2 (21)	rs2839372 (21)	rs11701058 (21)	21:47887791	0.07/0.28	0.01/0.11	0.33/0.95
<i>Cis</i>	SNORD14A (11)	rs2634462 (11)	rs6486334 (11)	11:17230389	0/0	0.07/0.62	0.04/0.59
<i>Cis</i>	TMEM149 (19)	rs807491 (19)	rs7254601 (19)	19:36234489	0/0.11	0.11/0.93	0.51/0.9
<i>Trans</i>	TMEM149 (19)	rs8106959 (19)	rs6926382 (6)	19:36234489	NA	0.84/0.99	NA
<i>Trans</i>	TMEM149 (19)	rs8106959 (19)	rs914940 (1)	19:36234489	NA	0.84/0.99	NA
<i>Trans</i>	TMEM149 (19)	rs8106959 (19)	rs2351458 (4)	19:36234489	NA	0.84/0.99	NA
<i>Trans</i>	TMEM149 (19)	rs8106959 (19)	rs6718480 (2)	19:36234489	NA	0.84/0.99	NA
<i>Trans</i>	TMEM149 (19)	rs8106959 (19)	rs1843357 (8)	19:36234489	NA	0.84/0.99	NA
<i>Trans</i>	TMEM149 (19)	rs8106959 (19)	rs9509428 (13)	19:36234489	NA	0.84/0.99	NA
<i>Cis</i>	VASP (19)	rs1264226 (19)	rs2276470 (19)	19:46033382	0.01/0.12	0.05/0.47	0.1/0.57

NA, not applicable because the SNPs are on different chromosomes.  
\* IncSeq variant is the most strongly associated additive variant with probe levels in *cis* ( $\pm$  1Mb probe start site).



**Figure 1 | Haplotype and linkage disequilibrium structure. a, b,** Haplotype and LD structure are shown at the ADK locus of two proposed epistatic SNPs from Hemani *et al.*<sup>1</sup> (a) and when adding a third SNP captured by sequencing in 450 Italian individuals (b). The two “epistatic” SNPs form all four of the possible haplotypes. When adding the third SNP no new haplotypes are formed at >2.4% frequency. Haplotypes were estimated using Haploview<sup>8</sup>.

of a technical explanation for the *cis-trans* interactions. Our results mean that the apparent epistasis reported by Hemani *et al.* is more likely to be due to moderate levels of LD between each of the two SNPs and a single causal allele rather than genuine epistasis.

Hemani *et al.* attempted to remove interaction effects driven by low levels of correlation with additive variants by removing pairs of SNPs with pairwise  $r^2 < 0.1$  and  $D'^2 < 0.1$  (Table 2). However, it is possible to have substantial multi-locus LD but no pairwise LD<sup>4</sup>. Fig. 1 provides an example of the haplotype structure for the ADK locus, where there is no LD between the two interacting SNPs, but the most associated variant from our study has moderate LD with both of the SNPs.

In summary, using whole-genome sequencing and independent data, we have provided an alternative explanation for the findings of Hemani *et al.*<sup>1</sup> and conclude that there remain few robust examples of epistasis in humans.

Methods

Gene expression profiles were captured using an Illumina HumanHT-12 v3.0 Bead-Chip array<sup>2</sup>. Whole-genome sequencing was performed at the Beijing Genomics Institute (Shenzhen, China) using the Illumina HiSeq 2000 (median read depth 7×). Reads were processed using GATK<sup>3</sup> before genotype recovery and refinement through within-sample imputation using BEAGLE<sup>6</sup>. Analysis of the 8 d.f. model and interaction term was performed using epiGPU<sup>3</sup>. To determine whether the observed interactions were driven by unaccounted for additive variants, we obtained the most strongly associated variant in *cis* (1 megabase  $\pm$  probe start site) using MACH2QTL<sup>7</sup>, generated a phenotype of residuals for each expression trait by regressing out the variant, and then repeated the epiGPU analysis using the adjusted trait.

Andrew R. Wood<sup>1</sup>, Marcus A. Tuke<sup>1</sup>, Mike A. Nalls<sup>2</sup>, Dena G. Hernandez<sup>2,3</sup>, Stefania Bandinelli<sup>4,5</sup>, Andrew B. Singleton<sup>2</sup>, David Melzer<sup>1</sup>, Luigi Ferrucci<sup>6</sup>, Timothy M. Frayling<sup>1</sup> & Michael N. Weedon<sup>1</sup>

<sup>1</sup>Institute of Biomedical and Clinical Sciences, University of Exeter Medical School, Exeter EX2 5DW, UK.

email: M.N.Weedon@exeter.ac.uk and T.M.Frayling@exeter.ac.uk

<sup>2</sup>Laboratory of Neurogenetics, National Institute of Aging, Bethesda, Maryland 20892, USA.

<sup>3</sup>Department of Molecular Neuroscience and Reta Lila Laboratories, Institute of Neurology, UCL, London WC1N 1PJ, UK.

<sup>4</sup>Tuscany Regional Health Agency, Florence, Italy, I.O.T. and Department of Medical and Surgical Critical Care, University of Florence, Florence, Italy.

<sup>5</sup>Geriatric Unit, Azienda Sanitaria di Firenze, Florence, Italy.

<sup>6</sup>Longitudinal Studies Section, Clinical Research Branch, Gerontology Research Center, National Institute on Aging, Baltimore, Maryland 21225, USA.

- Hemani, G., Theodoridis, A., Wei, W. & Haley, C. EpiGPU: exhaustive pairwise epistasis scans parallelized on consumer level graphics cards. *Bioinformatics* **27**, 1462–1465 (2011).
- Nielsen, D. M., Ehm, M. G., Zaykin, D. V. & Weir, B. S. Effect of two- and three-locus linkage disequilibrium on the power to detect marker/phenotype associations. *Genetics* **168**, 1029–1040 (2004).
- DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genet.* **43**, 491–498 (2011).
- Browning, B. L. & Yu, Z. Simultaneous genotype calling and haplotype phasing improves genotype accuracy and reduces false-positive associations for genome-wide association studies. *Am. J. Hum. Genet.* **85**, 847–861 (2009).
- Li, Y., Willer, C. J., Ding, J., Scheet, P. & Abecasis, G. R. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.* **34**, 816–834 (2010).
- Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).

**Author Contributions** A.R.W., T.M.F. and M.N.W. designed the study. A.R.W., M.A.T. and M.N.W. performed the bioinformatics analyses. M.A.N., D.G.H., S.B., A.B.S., D.M. and L.F. provided the InChianti study and expression data. A.R.W., T.M.F. and M.N.W. wrote the manuscript. All authors commented on the manuscript.

**Competing Financial Interests** Declared none.

doi:10.1038/nature13691

Received 10 March 2014; accepted 1 July 2014.

- Hemani, G. et al. Detection and replication of epistasis influencing transcription in humans. *Nature* **508**, 249–253 (2014).
- Wood, A. R. et al. Allelic heterogeneity and more detailed analyses of known loci explain additional phenotypic variation and reveal complex patterns of association. *Hum. Mol. Genet.* **20**, 4082–4092 (2011).

## Hemani et al. reply

REPLYING TO A. R. Wood et al. *Nature* **514**, <http://dx.doi.org/10.1038/nature13691> (2014)

We thank Wood et al. for their interesting observations in the accompanying Comment<sup>1</sup>, and although their proposed mechanism does not explain all our reported results, we acknowledge that alternative mechanisms could be behind the observation of epistatic signals. Although we replicate our results in large, independent samples, 19/30 of our reported

interactions (Table 1 in ref. 2), Wood et al.<sup>1</sup> do not replicate in the InChianti data set ( $n = 450$ ) at a type-I error rate of  $0.05/30 = 0.002$ , including none of our reported *cis-trans* interactions. Having insufficient data to replicate the discovery interactions makes it problematic to draw firm conclusions on the reported *cis-trans* effects.

**Table 1 | Meta-analysis of results from discovery and replication cohorts**

Cis/trans	Gene (chr)	SNP1 (chr)	SNP2 (chr)	IncSeq SNP from imputed data	Interaction $-\log(P \text{ value})$ (three studies)	Interaction $-\log(P \text{ value})$ (two studies)
Cis	ADK (10)	rs2395095 (10)	rs10824092 (10)	rs67594352	3.25	2.9
Cis	ATP13A1 (19)	rs4284750 (19)	rs873870 (19)	NA	NA	NA
Cis	C21ORF57 (21)	rs9978658 (21)	rs11701361 (21)	rs11702450	6.62	5.57
Cis	CSTB (21)	rs9979356 (21)	rs3761385 (21)	rs35285321	1.64	1.63
Cis	CTSC (11)	rs7930237 (11)	rs556895 (11)	rs56375235	10.53	7.88
Cis	FN3KRP (17)	rs898095 (17)	rs9892064 (17)	NA	NA	NA
Cis	GAA (17)	rs11150847 (17)	rs12602462 (17)	rs4889970	11.85	8.29
Cis	HNRPH1 (5)	rs6894268 (5)	rs4700810 (5)	rs10078796	10.82	4.91
Cis	LAX1 (1)	rs1891432 (1)	rs10900520 (1)	rs2185079	1.01	1
Cis	MBLN1 (3)	rs16864367 (3)	rs13079208 (3)	rs67903230	4.19	3.23
Trans	MBLN1 (3)	rs7710738 (5)	rs13069559 (3)	rs67903230	3.42	2.97
Trans	MBLN1 (3)	rs2030926 (6)	rs13069559 (3)	rs67903230	5.31	3.96
Trans	MBLN1 (3)	rs2614467 (14)	rs13069559 (3)	rs67903230	3.12	2.88
Trans	MBLN1 (3)	rs218671 (17)	rs13069559 (3)	rs67903230	4.85	2.84
Trans	MBLN1 (3)	rs11981513 (7)	rs13069559 (3)	rs67903230	6.49	5.75
Cis	MBP (18)	rs8092433 (18)	rs4890876 (18)	rs470929	4.08	3.27
Cis	NAPRT1 (8)	rs2123758 (8)	rs3889129 (8)	rs10093709	4.07	2.95
Cis	NCL (2)	rs7563453 (2)	rs4973397 (2)	rs13019380	3.48	3.24
Cis	PRMT2 (21)	rs2839372 (21)	rs11701058 (21)	rs4819255	15.80	12.16
Cis	SNORD14A (11)	rs2634462 (11)	rs6486334 (11)	rs2354863	5.01	3.66
Cis	TMEM149 (19)	rs807491 (19)	rs7254601 (19)	rs28656784	4.82	3.57
Trans	TMEM149 (19)	rs8106959 (19)	rs6926382 (6)	rs28656784	3.14	2.91
Trans	TMEM149 (19)	rs8106959 (19)	rs914940 (1)	rs28656784	3.47	3.12
Trans	TMEM149 (19)	rs8106959 (19)	rs2351458 (4)	rs28656784	4.77	4.01
Trans	TMEM149 (19)	rs8106959 (19)	rs6718480 (2)	rs28656784	4.86	3.69
Trans	TMEM149 (19)	rs8106959 (19)	rs1843357 (8)	rs28656784	3.34	3.14
Trans	TMEM149 (19)	rs8106959 (19)	rs9509428 (13)	rs28656784	3.06	2.73
Cis	VASP (19)	rs1264226 (19)	rs2276470 (19)	rs4803827	4.41	3.27

The analysis followed that of Wood et al.<sup>1</sup>. In each cohort the effect of the imputed IncSeq SNP was regressed against the probe levels and the residuals used as an adjusted phenotype. Interaction effects were estimated following Hemani et al.<sup>2</sup> and the results combined using Fisher's method (see Hemani et al.<sup>2</sup>) using results from all three data sets or just the two replication data sets. Two IncSeq SNPs were either not in the 1000 Genomes reference panel or did not pass imputation quality control. Remaining imputed IncSeq SNPs had imputation accuracy  $r^2 > 0.98$  in the Brisbane Systems Genetics Study (BSGS). Of the remaining 26, 24 had interaction  $P$  values  $< 0.05/26 = 1.9 \times 10^{-3}$ .



**Table 2 | Correlation coefficients are calculated between relative pairs in BSGS<sup>5</sup>**

ILMN_GENE	PROBE_ID	PP	PO	DZ	SIB	MZ	$h^2$	$d^2$
ADK	ILMN_2358626	0.01	0.14	0.12	0.09	0.38	0.41	0.12
ATP13A1	ILMN_2134224	-0.02	0.16	0.14	0.20	0.61	0.67	0.16
C21ORF57	ILMN_1795836	-0.02	0.15	0.17	0.23	0.47	0.51	0.08
CSTB	ILMN_1761797	-0.06	0.16	0.15	0.17	0.30	0.25	0.04
CTSC	ILMN_2242463	0.12	0.14	0.20	0.16	0.37	0.27	0.08
FN3KRP	ILMN_1652333	-0.07	0.17	0.14	0.21	0.43	0.31	0.11
GAA	ILMN_2410783	-0.05	0.16	0.14	0.13	0.39	0.39	0.06
HNRPH1	ILMN_2101920	0.01	0.15	0.12	0.13	0.24	0.17	0.05
LAX1	ILMN_1769782	-0.06	0.14	0.17	0.19	0.36	0.27	0.04
MBNL1	ILMN_2313158	0.02	0.18	0.16	0.18	0.42	0.18	0.11
NAPRT1	ILMN_1710752	-0.06	0.19	0.21	0.28	0.51	0.37	0.14
NCL	ILMN_2121437	-0.02	0.14	0.18	0.14	0.40	0.31	0.08
PRMT2	ILMN_1675038	-0.04	0.20	0.19	0.18	0.40	0.34	0.06
SNORD14A	ILMN_1799381	0.03	0.17	0.14	0.13	0.52	0.43	0.14
TMEM149	ILMN_1786426	0.06	0.27	0.23	0.17	0.49	0.41	0.09
VASP	ILMN_1743646	0.00	0.14	0.27	0.18	0.52	0.38	0.13

PP, parent–parent; PO, parent–offspring; DZ, dizygotic twins; SIB, sibling pairs not including DZ and MZ twins; MA, monozygotic twins. Estimates of additive ( $h^2$ ) and non-additive ( $d^2$ ) variance components estimated from pedigree data<sup>4</sup>. All probes are within the top 90th percentile of  $h^2$  estimates and the 95th percentile of  $d^2$  (from 17,994 probes).

TMEM149 and C21ORF57 are also known as IGFLR1 and YBEY, respectively.

Applying their method in our discovery and replication data sets<sup>2</sup> does not completely abrogate the statistical evidence for epistasis. Specifically, the meta-analysis of these results shows that weaker interaction effects remain for 24/26 epistasis pairs after correcting for effects of the IncSeq SNP (Table 1). For the remaining two pairs (at *CSTB* and *LAX1*) we cannot rule out a haplotype effect such as postulated by Wood *et al.*<sup>1</sup> and this may indeed be a more parsimonious explanation for these two pairs. Haplotype effects are known to be confounding factors in *cis*–*cis* interactions, as stated in Hemani *et al.*<sup>2</sup> The remaining results may remain significant owing to imperfect imputation of the IncSeq SNP (although imputation  $r^2$  is high), and we acknowledge that the presence of imperfectly tagged *cis* SNPs with large additive effects could lead to inflation of the *F*-statistic for epistatic interactions owing to violations of normality assumptions.

For 11 of the *cis*–*cis* pairs that were replicated by Wood *et al.*<sup>1</sup> there is evidence for additional *cis*-genetic variation to that explained by the IncSeq SNPs<sup>3</sup>. Hence the IncSeq SNPs are not the only (causal) variants in *cis* and therefore the additive effect of the IncSeq SNPs may contain additive effects of additional variants. Furthermore, these probes are within the 95th percentile of non-additive genetic variation estimated using a pedigree-based method that is completely orthogonal to SNP-based methods<sup>4</sup> (Table 2).

Finally, we note that we did not report that epistasis was widespread and pointed out that for gene expression additive genetic variation explains much more of the total genetic variation than non-additive variation<sup>2,4</sup>.

**Gibran Hemani<sup>1,2</sup>, Konstantin Shakhbazov<sup>1,2</sup>, Harm-Jan Westra<sup>3</sup>,  
Tonu Esko<sup>4,5,6</sup>, Anjali K. Henders<sup>7</sup>, Allan F. McRae<sup>1,2</sup>, Jian Yang<sup>1</sup>,**

**Greg Gibson<sup>8</sup>, Nicholas G. Martin<sup>7</sup>, Andres Metspalu<sup>4</sup>, Lude Franke<sup>3</sup>,  
Grant W. Montgomery<sup>7</sup>, Peter M. Visscher<sup>1,2</sup> & Joseph E. Powell<sup>1,2</sup>**

<sup>1</sup>Queensland Brain Institute, University of Queensland, Brisbane, Queensland 4072, Australia.

email: g.hemani@uq.edu.au

<sup>2</sup>University of Queensland Diamantina Institute, University of Queensland, Princess Alexandra Hospital, Brisbane, Queensland 4072, Australia.

<sup>3</sup>Department of Genetics, University Medical Center Groningen, University of Groningen, Hanzeplein 1, 9700 RB Groningen, The Netherlands.

<sup>4</sup>Estonian Genome Center, University of Tartu, Tartu, 51010, Estonia.

<sup>5</sup>Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts 02142, USA.

<sup>6</sup>Divisions of Endocrinology, Children's Hospital, Boston, Massachusetts 02115, USA.

<sup>7</sup>Queensland Institute of Medical Research, Brisbane, Queensland 4006, Australia.

<sup>8</sup>School of Biology and Centre for Integrative Genomics, Georgia Institute of Technology, Atlanta, Georgia 30332, USA.

1. Wood, A. R. *et al.* Another explanation for apparent epistasis. *Nature* **514**, <http://dx.doi.org/10.1038/nature13691> (2014).
2. Hemani, G. *et al.* Detection and replication of epistasis influencing transcription in humans. *Nature* **508**, 249–253 (2014).
3. Westra, H. J. *et al.* Systematic identification of *trans* eQTLs as putative drivers of known disease associations. *Nature Genet.* **45**, 1238–1243 (2013).
4. Powell, J. E. *et al.* Congruence of additive and non-additive effects on gene expression estimated from pedigree and SNP data. *PLoS Genet.* **9**, e1003502 (2013).
5. Powell, J. E. *et al.* The Brisbane Systems Genetics Study: genetical genomics meets complex trait genetics. *PLoS ONE* **7**, e35430 (2012).

doi:10.1038/nature13692

## GEOSCIENCE

# The plumbing of Greenland's ice

Observations of the water pressure in drilled boreholes and natural moulin on the Greenland Ice Sheet show how its underlying plumbing system controls ice motion during the course of the summer melt season. [SEE LETTER P.80](#)

PETER NIENOW

The interface between the base of an ice sheet or glacier and its underlying bed is of fundamental importance in controlling the speed at which the ice flows<sup>1–3</sup>. Of particular significance is how friction at the ice–bed interface is affected by the routing of meltwaters across the ice-sheet or glacier bed. On page 80 of this issue, Andrews and colleagues<sup>4</sup> make a fundamental advance in our understanding of the hydrology underlying the Greenland Ice Sheet and of how the evolution of the subglacial drainage system controls ice motion during the course of the summer melt season — when water is generated by melting of snow and ice at the ice-sheet surface. The authors demonstrate that, during the latter part of the melt season, variations in water pressure in subglacial channels control daily patterns of ice motion, but a longer term slowdown in ice flow is dependent on decreasing water pressures in areas away from the channels.

Studies of mountain glacier systems<sup>1,3,5</sup>, and more recently in Greenland<sup>6</sup>, have investigated how the subglacial drainage system evolves over the course of the melt season and how this evolution affects ice motion. At the onset of the melt season, meltwaters flow over the glacier or ice-sheet surface before draining into the ice through crevasses or moulins, large natural vertical pipes, which can route this meltwater rapidly to the glacier or ice-sheet bed<sup>7</sup> (Fig. 1). These initial meltwaters, on reaching the glacier bed, encounter a subglacial drainage system that is incapable of transporting the meltwater easily along the ice–bed interface. As a result, the water pressure in the subglacial drainage system increases, which decreases the friction at the ice–bed interface and the ice accelerates; in effect, the pressurized water is helping to partially float the overlying ice, enabling it to slide downhill more easily. However, as the volume of surface meltwaters routed to the glacier bed increases with warming summer temperatures, the water flowing across the bed starts to create, through the melting of ice at the ice–bed interface, subglacial channels that are more hydraulically efficient<sup>8,9</sup>. These channels enable the water to drain out of the glacier efficiently, thereby lowering the subglacial



JAMES BALOG/AURORA PHOTOS

**Figure 1 | A moulin on the Greenland Ice Sheet.** By monitoring water levels in large natural vertical pipes, which route water from melting surface snow and ice to the ice-sheet base, Andrews *et al.*<sup>4</sup> acquired evidence demonstrating the presence of hydraulically efficient subglacial channels fed by the surface meltwater.

water pressure, so that the glacier slows down due to the decreasing flotation effect.

To understand more fully how hydrology affects the dynamics of the Greenland Ice Sheet, Andrews and colleagues used a suite of methods to investigate the link between water pressure at the ice-sheet base and ice motion. They drilled holes, using a hot-water ‘drill’, through about 600 metres of ice to the ice-sheet bed and inserted pressure sensors into these boreholes to measure the subglacial water pressure at their base, while simultaneously monitoring ice motion at the surface using Global Positioning System (GPS) data. They also lowered pressure sensors into moulins located between about 0.3 and 1.6 kilometres from the boreholes to measure fluctuations in water level, and thus pressure, in the moulins.

Andrews *et al.* observed systematic differences between the water-pressure measurements in the moulins and the boreholes,

concluding that the moulins were connected to an efficient channelized component of the drainage system, whereas the boreholes monitored a hydraulically inefficient system unconnected to the channels. Water-pressure variations in the moulins (and thus channels) were positively correlated with daily patterns of ice motion, whereas borehole water pressures were anti-correlated. Water-pressure variations in the subglacial channels are therefore capable of affecting the friction at the ice–bed interface over a large enough area of the bed to enable the ice sheet to accelerate and decelerate on diurnal timescales. However, during the latter half of the melt season, ice motion gradually decreased, but mean moulin water levels (and thus channel pressure) remained relatively constant. By contrast, the water pressure in the boreholes decreased, implying that the longer-term seasonal slowdown was driven by changes in the unconnected subglacial



drainage system away from the large subglacial channels. These findings imply that, to understand the dynamic behaviour of the ice sheet, it is essential to understand the processes going on in areas distal to the subglacial channels as well as in the channels themselves.

Andrews and colleagues' findings corroborate many of the earlier detailed borehole observations from mountain glacier systems<sup>1,3,10–12</sup>, indicating that the processes controlling the interaction between the hydrology and dynamics of ice-sheet and smaller valley-glacier systems are similar. The authors' observations also demonstrate how difficult it is to drill directly into areas affected by pressure variations in subglacial channels, because these areas cover only a small fraction of the glacier bed, in contrast to the surrounding distributed drainage system<sup>3,8</sup>.

There remain considerable uncertainties regarding the processes linking the hydrology and dynamics of the Greenland Ice Sheet. The distance to which efficient subglacial channels

extend into the ice sheet during the melt season remains unclear; tests that use artificial tracers to track the speed at which water is routed from moulins to the ice-sheet margin indicate that efficient channels extend at least tens of kilometres into the ice sheet<sup>9</sup>, but will such channels extend further in a warming climate under enhanced surface melting? Furthermore, it is not clear whether the observations made are transferable to the rapidly moving tidewater glaciers — rivers of ice which move at approximately 1–10 km per year and are responsible for about half of the ice-mass loss from Greenland through the calving of large icebergs into the ocean<sup>13</sup>. The structure of these subglacial drainage systems, especially where the glaciers are flowing fastest as they near the ocean, is unknown, but is likely to be important in sustaining the high subglacial water pressures that enable the ice to slide so rapidly. Nevertheless, it is through further studies such as those by Andrews and colleagues that the complexities of the hydrological system lurking deep under

the thick ice of the Greenland Ice Sheet will be unravelled. ■

**Peter Nienow** is in the School of GeoSciences, University of Edinburgh, Edinburgh EH8 9XP, UK.

e-mail: [peter.nienow@ed.ac.uk](mailto:peter.nienow@ed.ac.uk)

1. Iken, A. & Bindenschadler, R. J. *Glaciol.* **32**, 101–119 (1986).
2. Alley, R. B. *et al.* *J. Geophys. Res.* **92**, 8921–8929 (1987).
3. Fountain, A. G. & Walder, J. S. *J. Rev. Geophys.* **36**, 299–328 (1998).
4. Andrews, L. C. *et al.* *Nature* **514**, 80–83 (2014).
5. Mair, D. *et al.* *J. Geophys. Res. Solid Earth* **107**, B8, 2175 (2002).
6. Bartholomew, I. *et al.* *Nature Geosci.* **3**, 408–411 (2010).
7. Catania, G. A. & Neumann, T. A. *Geophys. Res. Lett.* **37**, L02501 (2010).
8. Nienow, P. *et al.* *Earth Surf. Process.* **23**, 825–843 (1998).
9. Chandler, D. *et al.* *Nature Geosci.* **6**, 195–198 (2013).
10. Gordon, S. *et al.* *Hydrol. Process.* **12**, 105–133 (1998).
11. Hubbard, B. *et al.* *J. Glaciol.* **41**, 572–583 (1995).
12. Murray, T. & Clarke, G. K. C. *J. Geophys. Res. Solid Earth* **100**, 10231–10245 (1995).
13. Joughin, I. *et al.* *Science* **338**, 1172–1176 (2012).

## DNA REPAIR

# Making the cut

**Analysis of the first step in repairing double-stranded-DNA breaks reveals that the Mre11 enzyme makes a DNA nick at a point separate from the break ends, creating an entry site for further processing by exonuclease enzymes. SEE LETTER P.122**

LORRAINE S. SYMINGTON

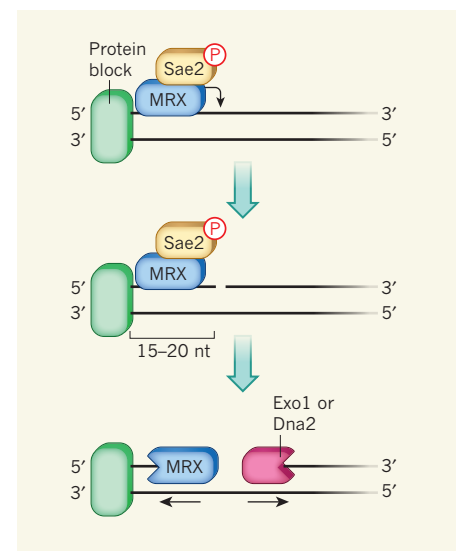
Double-stranded breaks in chromosomes are dangerous lesions that form when both strands of the DNA duplex are severed. Such breaks must be accurately repaired to preserve genome integrity — inaccurate or failed repair can result in chromosome rearrangements, loss of genetic information or even cell death. Indeed, faulty repair of double-stranded breaks is associated with infertility, developmental and immunological defects and predisposition to cancer. Cells repair breaks in an error-free manner through a mechanism called homologous recombination, which begins with removal of one of the two DNA strands at each broken end. On page 122 of this issue, Cannavo and Cejka<sup>1</sup> provide mechanistic insight into exactly how a three-protein complex, Mre11–Rad50–Xrs2, is involved in the initial stages of error-free DNA repair.

The first step in homologous recombination is the degradation of the 5' end of DNA on either side of a break to yield 3' single-stranded DNA (ssDNA) tails — a process called end resection. The Rad51 protein then binds to these tails and promotes exchange of genetic information with homologous sequences from the sister chromosome, leading to DNA

repair. Genetic studies<sup>2,3</sup> in the budding yeast *Saccharomyces cerevisiae* suggest a two-step mechanism for end resection. Initially, the evolutionarily conserved Mre11–Rad50–Xrs2 (MRX) enzymatic complex and the Sae2 protein clip off the 5'-terminated DNA strand, creating a short 3' overhang. Next, the overhang is rapidly lengthened by the Exo1 or Dna2 nuclease enzymes. These enzymes remove nucleotides from the strand being processed to generate an extensive tract of ssDNA.

Significant progress has been made in deciphering the mechanisms used by Exo1 and Dna2 for extensive resection<sup>4–6</sup>, but little is known about how MRX and Sae2 cooperate to initiate the process. Mre11 has exonuclease activity (it degrades DNA from the end of the strand), but an *in vitro* study<sup>7</sup> showed that the enzyme catalyses DNA degradation from 3' to 5', the opposite direction to that in which end resection occurs. Sae2 also reportedly shows nuclease activity *in vitro*<sup>8</sup>. A bidirectional model<sup>9</sup> proposes that MRX uses its 3'–5' exonuclease activity to proceed back to the double-stranded break from an internal nick in the DNA created by MRX and Sae2 (Fig. 1). However, the identity of the endonuclease that could create this internal nick has been unclear.

Cannavo and Cejka purified MRX and Sae2



**Figure 1 | Resection in the right direction.** Cannavo and Cejka<sup>1</sup> examined the role of the Mre11–Rad50–Xrs2 (MRX) enzyme complex and Sae2 protein in end resection, the first step in repairing double-stranded breaks in DNA. Using a protein block to mimic a natural DNA break, they found that MRX makes a nick in the DNA at 15 to 20 nucleotides (nt) from the 5' end of the break, a process that is promoted by phosphorylated (P) Sae2. The nick creates an entry site for Mre11 to proceed back to the break end in the 3' to 5' direction, and for the Exo1 or Dna2 enzymes to cleave 5' to 3', extending the resected end to leave a 3' single-stranded DNA tail.

and analysed the proteins' activities on model DNA substrates *in vitro*. They observed that MRX did indeed have 3'–5' exonuclease activity. But, in contrast to previous work<sup>8</sup>, they found no nuclease activity for Sae2 alone.

The authors incubated MRX and Sae2 with a linear double-stranded DNA substrate

in which one end was blocked by protein to mimic a double-stranded break. They observed an Sae2-dependent degradation product indicative of a nick internal to the protein-blocked DNA end. The researchers showed that this previously undocumented endonuclease activity was inherent to Mre11 — they repeated the experiment using a variant of the MRX complex containing a nuclease-defective version of Mre11, and observed no endonuclease activity.

Binding and hydrolysis of ATP molecules cause large conformational changes in the Rad50 subunit of the MRX complex, activating Mre11 nuclease activity<sup>10,11</sup>. Cannavo and Cejka observed MRX- and Sae2-dependent clipping activity only in the presence of ATP. This was eliminated using MRX mutants that were unable to bind or hydrolyse ATP. By adding protein blocks at both ends, the authors prevented the 3'–5' exonuclease action of Mre11 and, using labelling techniques, they then mapped the site at which the enzyme clipped the DNA to around 15 to 20 nucleotides from the break ends. Thus, the properties of the *in vitro* reaction reported by Cannavo and Cejka match the known requirements for homologous recombination *in vivo*<sup>2,3</sup>. Furthermore, this study explains how Mre11 promotes resection of the 5'-terminated strand.

One key question is how Sae2 regulates MRX endonuclease activity. Phosphate molecules, which can modify protein behaviour, are added to Sae2 by a protein-kinase enzyme through phosphorylation when cells enter the S phase of the cell cycle. Mutation of a serine amino-acid residue (serine 267) to an alanine residue that cannot be phosphorylated severely impairs Sae2 function *in vivo*<sup>12</sup>. Cannavo and Cejka found that this mutant protein rendered MRX unable to clip double-stranded DNA. In addition, they showed that dephosphorylation of Sae2 with a phosphatase enzyme decreased its activity. These data suggest that Sae2 acts as a phosphorylation-dependent switch to trigger MRX endonuclease activity.

Because MRX clipping is dependent on ATP, and a specific point mutation in the *RAD50* gene, like loss of Sae2, prevents DNA clipping<sup>13</sup>, it is likely that Sae2 acts through Rad50 to activate Mre11. Cannavo and Cejka detected a physical interaction between Sae2 and MRX, but found that, instead of the anticipated interaction with Rad50, only the Mre11 and Xrs2 subunits interacted with Sae2.

The CtIP protein is considered to be the equivalent of Sae2 in vertebrate cells and plays a crucial part in end resection. Two papers published earlier this year reported that CtIP has nuclease activity<sup>14,15</sup>. However, the protein's active site is not in its evolutionarily conserved carboxy-terminal domain, which, as the current study shows, is required in Sae2 to stimulate the latent MRX endonuclease. This raises the question of whether CtIP has further functions and can process DNA

independently of the vertebrate form of MRX.

Cannavo and Cejka's study provides mechanistic insight into how MRX and Sae2 initiate end processing, but leads to more questions. For example, how does MRX recognize a protein-blocked end? How does a blocked end trigger cleavage of the 5' strand? Finally, there is the question of how Sae2 phosphorylation coordinates with the ATPase activity of MRX to activate Mre11 endonuclease activity. ■

**Lorraine S. Symington** is in the Department of Microbiology and Immunology, Columbia University Medical Center, New York, New York 10032, USA.  
e-mail: lss5@columbia.edu

1. Cannavo, E. & Cejka, P. *Nature* **514**, 122–125 (2014).
2. Mimitou, E. P. & Symington, L. S. *Nature* **455**, 770–774 (2008).

#### SOLAR SYSTEM

## Not just a storm in a teacup

**The cloud that emerged above the south pole of Saturn's moon Titan in 2012 has been found to consist of hydrogen cyanide particles. This unexpected result prompts fresh thinking about the atmosphere of this satellite. SEE LETTER P.65**

CAITLIN A. GRIFFITH

**I**n May 2012, a large spinning cloud appeared over the southern pole of Saturn's moon Titan, where it persists still<sup>1</sup> (Fig. 1). The composition of the cloud has eluded identification until now. On page 65 of this issue, de Kok *et al.*<sup>2</sup> provide strong evidence for an unexpected answer: the cloud is made of hydrogen cyanide (HCN) ice particles. This result is difficult to refute because two spectral features indicate the presence of HCN, rather than of a HCN polymer<sup>3</sup>, and the cloud's mass is consistent with that predicted for a cloud composed of HCN. The only problem is that the cloud, at an altitude of 300 kilometres, is not where it is supposed to be.

HCN is expected to condense in Titan's atmosphere at an altitude of 80 km; indeed, a nearly imperceptible tropical haze layer at this height matches the anticipated effects of HCN condensation<sup>4</sup>. By contrast, as discussed by de Kok and colleagues, the south polar HCN cloud emerged in the satellite's southern polar winter and resides in a region where, only three months earlier, the Cassini spacecraft's infrared spectrometer measured a temperature of 170 kelvin, which is 45 K too warm for HCN to condense<sup>5</sup>.

The next temperature measurement by Cassini will occur in 2015. Perhaps these

3. Zhu, Z., Chung, W.-H., Shim, E. Y., Lee, S. E. & Ira, G. *Cell* **134**, 981–994 (2008).
4. Cannavo, E., Cejka, P. & Kowalczykowski, S. C. *Proc. Natl Acad. Sci. USA* **110**, E1661–E1668 (2013).
5. Niu, H. *et al.* *Nature* **467**, 108–111 (2010).
6. Cejka, P. *et al.* *Nature* **467**, 112–116 (2010).
7. Paull, T. T. & Gellert, M. *Mol. Cell* **1**, 969–979 (1998).
8. Lengsfeld, B. M., Rattray, A. J., Bhaskara, V., Ghirlando, R. & Paull, T. T. *Mol. Cell* **28**, 638–651 (2007).
9. Garcia, V., Phelps, S. E. L., Gray, S. & Neale, M. J. *Nature* **479**, 241–244 (2011).
10. Lammens, K. *et al.* *Cell* **145**, 54–66 (2011).
11. Lim, H. S., Kim, J. S., Park, Y. B., Gwon, G. H. & Cho, Y. *Genes Dev.* **25**, 1091–1104 (2011).
12. Huertas, P., Cortés-Ledesma, F., Sartori, A. A., Aguilera, A. & Jackson, S. P. *Nature* **455**, 689–692 (2008).
13. Alani, E., Padmore, R. & Kleckner, N. *Cell* **61**, 419–436 (1990).
14. Wang, H. *et al.* *Mol. Cell* **54**, 1012–1021 (2014).
15. Makharashvili, N. *et al.* *Mol. Cell* **54**, 1022–1033 (2014).

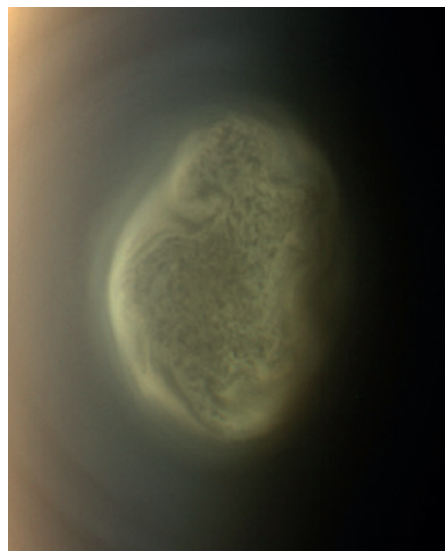
This article was published online on 17 September 2014.

data will reveal that Titan's atmosphere is appropriately cool at an altitude of 300 km, arming theorists with enough information to understand the complex conditions of the polar winter. Barring this inconvenience with the temperature, HCN clouds at high polar altitudes can form by processes that are typical of Earth's atmosphere, a point that becomes apparent when considering the broader context of Titan's atmospheric chemistry and dynamics.

Titan's atmospheric composition resembles certain models of early Earth, before oxygen was a significant component of our atmosphere and when methane (CH<sub>4</sub>) may have carried much of the atmospheric carbon<sup>6</sup>. Titan's two most abundant constituents, nitrogen (N<sub>2</sub>) and methane, control the atmospheric make-up. These molecules are broken apart in the upper atmosphere (at roughly 1,000 km altitude) by solar ultraviolet radiation in a process called photolysis, thereby yielding reactive radicals, which initiate the production of complex organic molecules. The main nitrogen-containing molecule produced, HCN, regulates the production of nitrogen species.

The photochemically produced molecules mix down to the lower atmosphere while chemically mingling to form new molecules. Eventually, they settle on the moon's surface





**Figure 1 | Titan's south polar cloud.** The cloud rises above the winter darkness of Titan's southern pole and catches the glint of sunlight. De Kok *et al.*<sup>2</sup> find that this spinning cloud, which completes an orbit in 9 hours — much faster than Titan's 15-day orbit — is composed of hydrogen cyanide. This orbital period signals the formation of a polar vortex as the southern pole enters into winter darkness<sup>14</sup>. The image, taken by the Cassini spacecraft's Imaging Subsystem, is centred near the south pole and extends over a distance of about 1,000 kilometres.

in organic lakes and puddles, with a 'soup base' of methane and ethane ( $C_2H_6$ ) that resembles natural gas<sup>7,8</sup>. As these nitrogen-bearing molecules diffuse downward, they are steered towards Titan's winter pole by atmospheric circulation<sup>9</sup>. Here they enter cooler regions and condense into clouds at different atmospheric layers, depending on their different thermodynamic properties.

Two distinct kinds of cloud cap the winter pole, both identified from measurements made by Cassini's Visual and Infrared Mapping Spectrometer: one at 55 km altitude, which is consistent with a  $C_2H_6$  composition<sup>10</sup>, and one at 300 km, found by de Kok and colleagues to be made of HCN. These stacked clouds, composed of the most abundant nitrogen and carbon photochemical species, track seasonally with the winter pole. An HCN cloud was previously identified<sup>8</sup> above Titan's north pole at the end of the northern winter, and the  $C_2H_6$  cloud has vanished from this region in preparation for its southerly migration for the winter<sup>11</sup>.

The winter pole on Titan is a peculiar place. Here the atmosphere radiatively cools during winter's darkness, triggering a suite of dynamical atmospheric responses. As discussed by de Kok and colleagues, the polar temperature is regulated by the intertwined effects of atmospheric chemistry, radiation and dynamics, which control the atmosphere's absorption and emission of radiation, and the compression, expansion and mixing

of Titan's gases. A potential explanation for the polar HCN clouds is that they form from a process known as open-cell convection, in which cool, dense air sinks and warms slightly while the surrounding air rises and cools, thereby forming clouds (Fig. 1). The cool polar atmosphere also contrasts with the warmer lower latitudes. The resulting decrease in temperature with increasing latitude affects the atmosphere's pressure structure, which, when combined with Titan's spin, implies circumpolar winds that become more strongly westerly with altitude<sup>12</sup>. Titan's atmospheric polar vortex, witnessed by the rapid spin of the south polar HCN cloud, isolates the polar air from the rest of the atmosphere, allowing the pole to cool further. The presence of clouds spinning in a vortex can thus naturally emerge at a winter pole.

However, the detailed operations of Titan's winter pole, such as the seasonal evolution of the chemistry and temperature at 300 km altitude, are complicated and far from understood<sup>12</sup>. Better grasped is Earth's polar atmosphere, where the winter polar vortex is a repository of unique composition and clouds. The polar chemistry evolves from winter to spring with the production and loss of many molecular species that ultimately control the polar ozone abundances.

Laboratory simulations suggest that the photochemistry in Titan's atmosphere produces amino acids and nucleotide bases<sup>13</sup>. How far the chemistry evolves in Titan's upper atmosphere is unclear, and would probably require detailed *in situ* sampling of the upper

atmosphere. However, further understanding of Titan's organic chemistry will entail studies of the abundance, phase and particulate composition of the main nitrogen photochemical product, which affects the overall nitrogen chemistry. The presence of an HCN maelstrom opens investigations into a new avenue of planetary-satellite organic chemistry — that of a cold and dark polar vortex stocked with nitrogen and methane photolysis products, which are typical of Titan and, perhaps, of early Earth. ■

**Caitlin A. Griffith** is in the Department of Planetary Sciences, University of Arizona, Tucson, Arizona 8521-0092, USA.  
e-mail: griffith@lpl.arizona.edu

- West, R. A. *et al.* Abstr. 305.03 (AAS/Division for Planetary Sciences Meeting Abstracts, Vol. 45, 2013).
- de Kok, R. J. *et al.* *Nature* **514**, 65–67 (2014).
- Cruikshank, D. P. *et al.* *Icarus* **94**, 345–353 (1991).
- Lavvas, P., Griffith, C. A. & Yelle, R. V. *Icarus* **215**, 732–750 (2011).
- Vinatar, S. *et al.* *Icarus* (submitted).
- Trainer, M. G. *et al.* *Proc. Natl Acad. Sci. USA* **103**, 18035–18042 (2006).
- Yung, Y. L., Allen, M. & Pinto, J. P. *Astrophys. J. Suppl.* **55**, 465–506 (1984).
- Clark, R. N. *et al.* *J. Geophys. Res. Planets* **115**, E10005 (2010).
- Rannou, P., Lebonnois, S., Hourdin, F. & Luz, D. *Adv. Space Res.* **36**, 2194–2198 (2005).
- Griffith, C. A. *et al.* *Science* **313**, 1620–1622 (2006).
- Le Mouélic, S. *et al.* *Planet. Space Sci.* **60**, 86–92 (2012).
- Flasar, F. M. & Achterberg, R. K. *Phil. Trans. R. Soc. A* **367**, 649–664 (2009).
- Hörst, S. M. *et al.* *Astrobiology* **12**, 809–817 (2012).
- <http://photojournal.jpl.nasa.gov/catalog/PIA14920>

#### MEDICAL RESEARCH

## Ebola therapy protects severely ill monkeys

**A blend of three monoclonal antibodies has completely protected monkeys against a lethal dose of Ebola virus. Unlike other post-infection therapies, the treatment works even at advanced stages of the disease. [SEE ARTICLE P.47](#)**

THOMAS W. GEISBERT

**T**he filoviruses known as Ebola virus and Marburg virus are among the most deadly of pathogens, with fatality rates of up to 90% (ref. 1). Early this year, a new strain of the Zaire species of Ebola virus emerged<sup>2</sup> in the West African country of Guinea and quickly spread to Liberia, Sierra Leone and Nigeria. The outbreak persists despite the best efforts of local and international authorities, and is now the largest filovirus outbreak on record, with no end in sight. There are no licensed vaccines or

post-exposure treatments against Ebola, so moving the most promising interventions forward is a matter of utmost urgency. On page 47 of this issue, Qiu *et al.*<sup>3</sup> report that rhesus monkeys can be completely protected from lethal Ebola infection using ZMapp — a blend of three monoclonal antibodies. Crucially, the treatment protected monkeys even when it was administered as late as 5 days after exposure to the virus, at a time when the animals were severely ill.

Since the discovery of Ebola virus (Fig. 1) in 1976, researchers have been actively developing treatments to combat infection. Studies

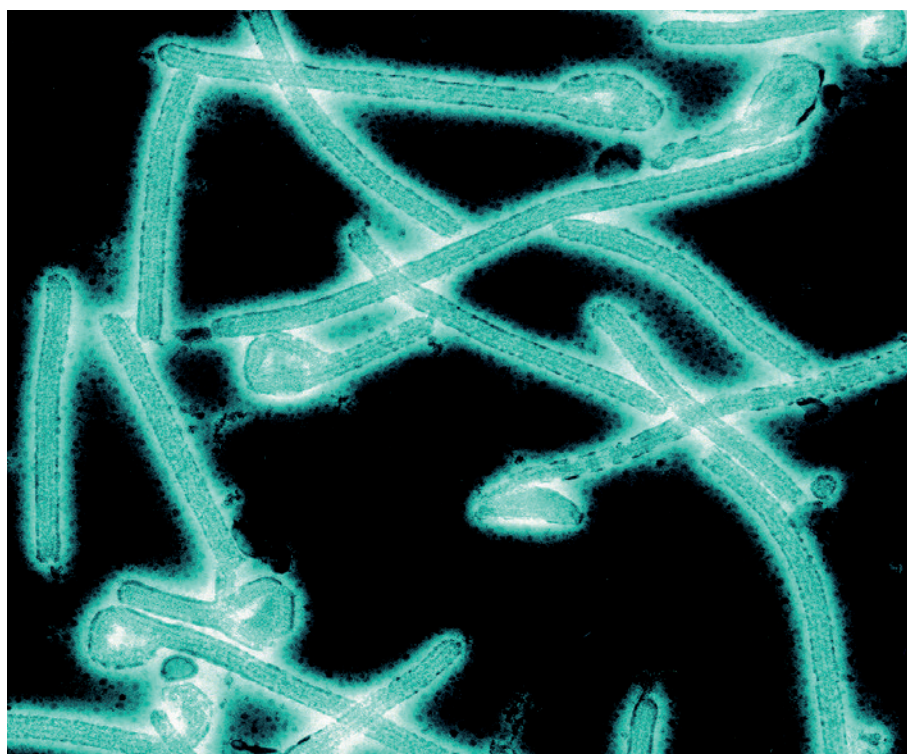


Figure 1 | The Ebola virus.

over the past decade have found that modulators of blood coagulation<sup>4,5</sup>, an antisense oligonucleotide called AVI-6002 (ref. 6) and a vaccine<sup>7</sup> based on vesicular stomatitis virus (VSV) all afforded partial protection of monkeys against Ebola when administered within an hour of virus exposure. The VSV-based vaccine was used in 2009 to treat a laboratory worker in Germany shortly after she was accidentally pricked with a needle possibly contaminated by an Ebola-infected animal<sup>8</sup>. The worker survived, but it is unclear whether this was because she had not been exposed to Ebola or because the vaccine protected her.

Subsequent advances have been made in developing treatments that can completely protect monkeys against Ebola. These include small ‘interfering’ RNAs (known as TKM-Ebola<sup>9</sup>) and various combinations of antibodies<sup>10–12</sup>. But these treatments need to be administered within 2 days of exposure to the virus. So although these approaches were highly important and can be used to treat known exposures, the need for treatments that protect at later times after infection was paramount.

Further development and improvement of the antibody-based strategies led to a cocktail of monoclonal antibodies<sup>13</sup> that protected 43% of monkeys when given as late as 5 days after Ebola exposure — a time at which the clinical signs of disease are apparent. Another therapy that combines monoclonal antibodies with interferon- $\alpha$  (a protein that stimulates an antiviral response) provides almost complete protection of macaques when given 3 days after exposure<sup>14</sup>, at which point

the virus can be detected but clinical signs are only just beginning to be seen in some animals.

Qiu *et al.* now report ZMapp, an antibody therapy that does not require interferon- $\alpha$ , and which was developed by two collaborating teams of researchers who had worked on some of the previously reported antibody treatments. ZMapp was made by testing different combinations of chimaeric monoclonal antibodies (in which fragments of human antibodies are attached to antibody fragments from mice). The optimal formulation contains two antibodies from a previously reported blend<sup>14</sup> and a third from a different cocktail<sup>13</sup>.

To test the therapy, Qiu *et al.* administered a lethal dose of Ebola virus to three groups of six animals, and then treated them with three doses of ZMapp. The first group received therapy at 3, 6 and 9 days post-infection; the second group at 4, 7 and 10 days; and the third group at 5, 8 and 11 days. Remarkably, all the animals survived, and were found to have undetectable viral loads by 21 days after infection. It should be noted that the authors used the Kikwit variant of the virus in these experiments, because the Guinean strain from the current West African outbreak was not available in time for this part of their study. However, they went on to show that ZMapp inhibits replication of the Guinean strain in cell culture.

The development of ZMapp and its success in treating monkeys at an advanced stage of Ebola infection is a monumental achievement. On this basis, the treatment has been used in the current Ebola outbreak to treat several patients on compassionate grounds<sup>15</sup>. Of these, two US health-care workers have recovered

— but whether ZMapp had any effect is unknown, because at the time of writing, about 45% of patients in this outbreak survive without treatment<sup>16</sup>. As of 26 August, two other patients treated with ZMapp have not survived, but this might be because the treatment was initiated too late in the course of the disease.

The diversity of strains and species of Ebola and Marburg viruses remain an obstacle for all candidate treatments. Lethal disease in humans is caused by three different species of Ebola virus (Sudan, Bundibugyo and Zaire) and two genetically distinct lineages of Marburg virus. Treatments that protect against one species of Ebola — Zaire, in the case of ZMapp — will probably not protect against a different species of the virus, and might not protect against a different strain within a species.

Although the need for treatments for filovirus infections is unquestionable, the most effective way to manage and control future outbreaks might be through preventive vaccines, some of which have been tailored to protect against multiple species and strains. During outbreaks, single-injection vaccines are needed to ensure rapid use and protection. At least five preventive vaccines have been shown to completely protect monkeys against Ebola and Marburg infection<sup>17</sup>. But only VSV-based vaccines have been reported to completely protect monkeys against Ebola (Zaire) virus after a single injection<sup>18</sup> — notably, the wild-type virus, rather than a cultured variant that has also been used in research, and which produces slower disease progression in macaques.

Antibody therapies and several other strategies mentioned here should ultimately be included in an arsenal of interventions for controlling future Ebola outbreaks. Although ZMapp in particular has been administered for compassionate use, the next crucial step will be to formally assess its safety and effectiveness. Testing the latter is clearly difficult, because intentional infection of human subjects in clinical trials is not possible. US regulations, however, could allow the treatment to be licensed for widespread use on the basis of safety testing in humans and efficacy testing in animals. In the long run, the manufacture of ZMapp could require investment in infrastructure for making monoclonal antibodies at an industrial scale — assuming that funding is available to pay the production costs. ■

**Thomas W. Geisbert** is at the University of Texas Medical Branch at Galveston, Galveston National Laboratory, Galveston, Texas 77550-0610, USA.  
e-mail: [twgeisbe@utmb.edu](mailto:twgeisbe@utmb.edu)

1. Feldmann, H., Sanchez, A. & Geisbert, T. W. in *Fields Virology* 6th edn (eds Knipe, D. M. & Howley, P. M.) 923–956 (Lippincott Williams & Wilkins, 2013).
2. Baize, S. *et al.* *N. Engl. J. Med.* <http://dx.doi.org/10.1056/NEJMoa1404505> (2014).
3. Qiu, X. *et al.* *Nature* **514**, 47–53 (2014).

THOMAS W. GEISBERT



4. Geisbert, T. W. *et al.* *Lancet* **362**, 1953–1958 (2003).
5. Hensley, L. E. *et al.* *J. Infect. Dis.* **196** (Suppl. 2), S390–S399 (2007).
6. Warren, T. K. *et al.* *Nature Med.* **16**, 991–994 (2010).
7. Feldmann, H. *et al.* *PLoS Pathog.* **3**, e2 (2007).
8. Günther, S. *et al.* *J. Infect. Dis.* **204** (Suppl. 3), S785–S790 (2011).
9. Geisbert, T. W. *et al.* *Lancet* **375**, 1896–1905 (2010).
10. Dye, J. M. *et al.* *Proc. Natl Acad. Sci. USA* **109**, 5034–5039 (2012).
11. Olinger, G. G. Jr *et al.* *Proc. Natl Acad. Sci. USA* **109**, 18030–18035 (2012).
12. Qiu, X. *et al.* *Sci. Transl. Med.* **4**, 138ra81 (2012).
13. Pettitt, J. *et al.* *Sci. Transl. Med.* **5**, 199ra113 (2013).
14. Qiu, X. *et al.* *Sci. Transl. Med.* **5**, 207ra143 (2013).
15. Check Hayden, E. & Reardon, S. *Nature* <http://dx.doi.org/10.1038/nature.2014.15698> (2014).
16. [www.afro.who.int/en/clusters-a-programmes/dpc/epidemic-a-pandemic-alert-and-response/outbreak-news/4260-ebola-virus-disease-west-africa-22-august-2014.html](http://www.afro.who.int/en/clusters-a-programmes/dpc/epidemic-a-pandemic-alert-and-response/outbreak-news/4260-ebola-virus-disease-west-africa-22-august-2014.html)
17. Marzi, A. & Feldmann, H. *Expert Rev. Vaccines* **13**, 521–531 (2014).
18. Reported at the Filovirus Medical Countermeasure Workshop, 22–23 August 2013; <https://respond.niaid.nih.gov/conferences/mcmworkshop/Pages/default.aspx>

This article was published online on 29 August 2014.

## ASTRONOMY

# The age of the quasars

**An infrared census of accreting supermassive black holes across a wide range of cosmic times indicates that the canonical understanding of how these luminous objects form and evolve may need to be adjusted.**

DANIEL MORTLOCK

Ask an astronomer when quasars were at their peak and they will probably tell you it was about 10 billion years ago, when the Universe was about one-third of its current size<sup>1,2</sup>. Before then, the quasar population was still growing along with other large structures in the young Universe; there has since been a steady decrease in quasar numbers. However, in a paper published in *The Astrophysical Journal*, Vardanyan *et al.*<sup>3</sup> present results suggesting that this widely accepted picture may not be correct — or at least that it does not tell the whole story.

That story started in 1963 with the discovery<sup>4,5</sup> of a new type of astronomical object, referred to variously as quasi-stellar objects or quasars, the name that is generically used today. Their physical nature was initially unknown, but it was gradually deduced<sup>6</sup> that a quasar is a glowing disk of hot, dense material that can form around the supermassive black hole at the centre of a large galaxy, often the result of a collision with a second galaxy. Although such accretion disks are ‘only’ about the size of the Solar System, they can outshine all the stars in the host galaxy by a factor of a thousand or so. Quasars can hence be seen comparatively easily at great distances, which makes it possible to trace their evolution back to the first billion years after the Big Bang.

More than a million quasars have been catalogued in the 50 years since their discovery. Although this is more than enough for most demographic studies of astronomical objects, it is difficult to obtain a representative sample of quasars that spans a wide range of distances from Earth, and hence cosmic look-back times. It is also challenging to properly account for all the energy output of a quasar, because some

of the ultraviolet light that is emitted from the accretion disk is absorbed by dust in the host galaxy and re-radiated at much longer, infrared wavelengths. Most surveys of the quasar population have been undertaken using observations made at optical or near-infrared wavelengths (between about 0.2 and 2 micrometres), and it is these types of measurement that have provided the strongest evidence that quasar numbers peaked fairly sharply 10 billion years ago.

Vardanyan and colleagues studied a comparatively small sample of 10,000 quasars that were initially identified using optical data from the Sloan Digital Sky Survey. But, crucially, the authors had access to longer wavelength measurements (at about 8  $\mu\text{m}$ ) of the same objects from the Wide-Field Infrared Survey Explorer (WISE) satellite. They were thus able to get a more complete census of the quasars’ energy output and, after correcting for the various complicated observational selection effects that inevitably make such studies so difficult, found some striking results. They confirmed the steady decrease in the quasar population over the past 10 billion years but, rather than the expected drop at cosmic times before 3 billion years, they found a ‘plateau’ in the quasars’ energy output back to a little over a billion years after the Big Bang (Fig. 1). The authors were unable to probe any earlier than this, and one of their conclusions was that extending these sorts of measurements to earlier times is the best way to explore this issue further.

These results are not unprecedented — there have been several similar previous claims<sup>7,8</sup> that the canonical understanding of the quasar population from optical data was incomplete. However, the scale and quality of the WISE data are superior to any previously available. The findings demand serious attention, both



## 50 Years Ago

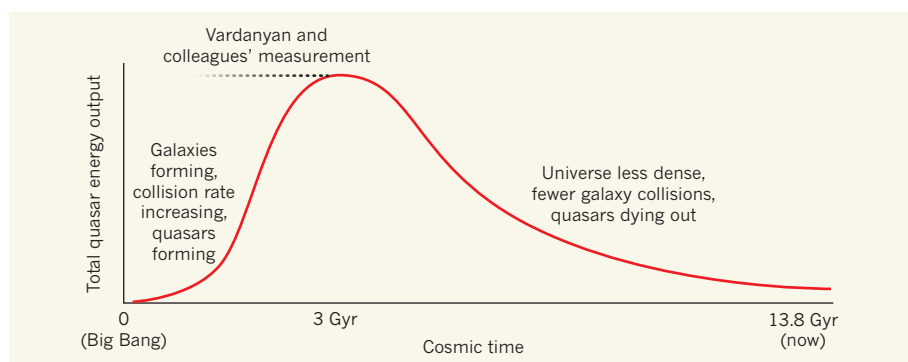
There are many puzzles about left handedness. Left handers are often less consistent in using the left hand than right handers in using the right; the incidence of left handedness is raised in many pathological groups and yet left handers may be of high intelligence; several different solutions have been offered to the problem of which cerebral hemisphere leads in speech functions in left handers; left handers seem to be more likely to recover from aphasia than right handers. It is the purpose of this article to describe a model of the inheritance of handedness and cerebral dominance which, together with a hypothesis about the direction of shifts of dominance, might account for many puzzling facts.

From *Nature* 3 October 1964

## 100 Years Ago

Every important town in Great Britain has established at least one great technical college at large cost in building and apparatus, with staffs of professors and teachers (always badly paid), and it is found that for their first two years the students have to be kept at great cost to the country learning those simple principles of science which they ought to have learnt at school. It is found that they are not only ignorant, but they have none of the habits of thought and scientific method which school laboratory work induces. The clever ones, if they leave school at seventeen, recover from the effects of a school education which prepared men only for being lawyers or clergyman; but the average man finds that he has been prepared only to be a hewer of wood and a drawer of water to the real engineer.

From *Nature* 1 October 1914



**Figure 1 | The energy output of quasars over cosmic time.** In the standard picture of quasar evolution, from the Big Bang to the present day, the total energy output of quasars increases to a peak value some 3 billion years (Gyr) after the Big Bang as galaxies form, collide and trigger the activation of quasars. This output then declines steadily as the accelerating expansion of the Universe results in a decrease in the number of galaxy collisions. Vardanyan *et al.*<sup>3</sup> found a surprising 'plateau' (dashed line) from about 1 billion to 3 billion years in the quasars' energy output.

in terms of subjecting them to further scrutiny and exploring their implications for quasar formation if the simplest interpretation — that large numbers of high-luminosity quasars were in place just a billion years after the Big Bang — is indeed correct.

The most exciting potential implication of Vardanyan and colleagues' study is that we need to adjust our understanding of the quasar population, especially how the early quasars formed. Most current models are based on the idea that galaxy collisions trigger quasar activation, so the number of quasars should rise sharply as galaxies form, grow and collide in the early Universe. The authors' results suggest that this link is not so strong, and that the most luminous quasars in particular form more rapidly than astronomers might suspect

using simple models of black-hole accretion and galaxy collisions.

The word 'suspect' is appropriate here, because this sort of science really is like detective work, in which indirect clues must be combined with inspired deduction to reach any interesting conclusions. It is remarkable that it is possible to make any kind of inference about black holes that are billions of light years away and have long since ceased to exist as quasars. One ambiguity is that the infrared light being used to assess the quasars' energy output could come from other sources, because any mechanism that heated whatever dust was present in the host galaxy would contribute to this signal. Also problematic is that various corrections to the inferred output of the quasars have to account for the

expansion of the Universe: the light seen at any given wavelength here and now has, since its emission, been redshifted by an amount that depends on how distant the source is, and hence how far back in time astronomers are seeing it. Perhaps the most uncertain aspect of all attempts to measure the evolution of the quasar population is deciding how best to account for this effect and how to test whether it has been done correctly. The approach taken by Vardanyan *et al.* is reasonable, but it is easy to imagine future data that would allow these corrections to be improved.

'More data' is something of a mantra in astronomy. Technological developments such as WISE have been one of the main drivers of discovery for the past century, and probably will continue to be in the future. We already have exciting projects such as the Large Synoptic Survey Telescope and the Square Kilometre Array just a few tantalizing years away, and both should tell us a great deal more about the age of the quasars. ■

**Daniel Mortlock** is in the Departments of Physics and of Mathematics, Imperial College London, London SW7 2AZ, UK.  
e-mail: d.mortlock@imperial.ac.uk

1. Fan, X. *et al. Astron. J.* **122**, 2833 (2001).
2. Richards, G. T. *et al. Astron. J.* **131**, 2766 (2006).
3. Vardanyan, V., Weedman, D. & Sargsyan, L. *Astrophys. J.* **790**, 88 (2014).
4. Hazard, C., Mackey, M. B. & Shimmins, A. J. *Nature* **197**, 1037–1039 (1963).
5. Schmidt, M. *Nature* **197**, 1040 (1963).
6. Rees, M. J. *Ann. Rev. Astron. Astrophys.* **22**, 471–506 (1984).
7. Casey, C. M. *et al. Astrophys. J.* **761**, 139 (2012).
8. Carilli, C. L. *et al. Astrophys. J.* **763**, 120 (2013).

that cause atopic dermatitis, psoriasis and acne and that are responsible for the inability of chronic wounds to heal. Yet the vast majority of our resident skin microorganisms are non-pathogenic, and many of these probably contribute to maintaining health. Indeed, earlier work from the group reporting the present study showed that, in healthy individuals, physiologically comparable body sites harbour similar bacterial and fungal communities<sup>2,3</sup>, and that shifts in skin communities are associated with development and immune status<sup>4,5</sup>. These results demonstrate that, instead of merely sampling the random bacteria in our environment with which our bodies interact, the skin can differentially select for specific populations.

The researchers have now moved beyond the question of which microbes are present on the skin to assessing what they might be doing. In this study, the authors sampled 15 healthy individuals at 18 sites and sequenced the metagenome — the collection of genomes in an environment — from each sample (Fig. 1). The use of metagenomic sequencing combined with innovative bioinformatic analyses

## MICROBIOLOGY

# An integrated view of the skin microbiome

**An analysis of the combined genomes of microorganisms inhabiting human skin demonstrates how these communities vary between individuals and across body sites, and paves the way to understanding their functions. SEE ARTICLE P.59**

PATRICK D. SCHLOSS

The growing interest in the human body's resident communities of microorganisms has paralleled a growing interest in probiotics and the emerging concept that foods can shape the composition of our gut microbiota and thus our health. At the same time, fuelled by fears of viruses and bacterial pathogens, hand sanitizers have become ubiquitous. The disconnect between protecting the balance of the  $10^{14}$  bacteria that reside within us

and destroying the  $10^{10}$  bacteria that live on us is jarring. However, our knowledge of the skin microbiota pales in comparison with that of our gut microbiota. Seeking to fill these gaps, on page 59 of this issue, Oh *et al.*<sup>1</sup> present an analysis of the genetic content of the bacteria, viruses and other microorganisms that live on human skin.

There is cause to distrust some of the microbes living on our skin — opportunistic pathogens such as *Staphylococcus aureus* reside there, as do the mixture of microbes



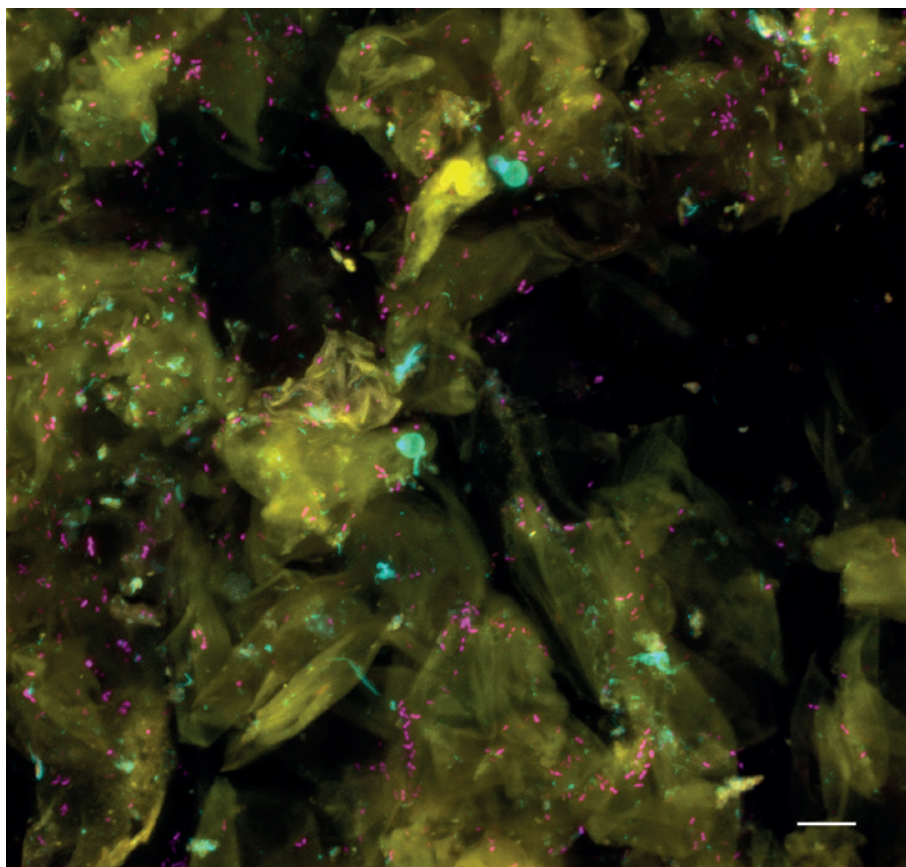
enabled them to obtain a more comprehensive taxonomic and genetic characterization of skin microbiota than has been previously attempted. Their results included not only bacteria, but also DNA viruses and microbial eukaryotes (nucleated organisms, such as protists and fungi).

This comprehensive survey revealed that each individual has a unique skin microbiota. The authors used these data to create a classifier, using a random-forest algorithm, that could differentiate between the 15 individuals on the basis of the skin metagenome, with a 19.3% error rate. When the authors attempted to classify the individuals using the bacterial, eukaryotic and viral data separately, the error rates were higher. Interestingly, it was not the dominant organisms, but the low-abundance organisms, that differentiated people. For example, the presence of Merkel cell polyomavirus, *Gardnerella vaginalis* and *Streptococcus pyogenes* were among the key features that could be used to differentiate between the individuals.

Among the more abundant bacterial populations, the researchers identified numerous strains of *Propionibacterium acnes* and *Staphylococcus epidermidis*. Investigating the spatial and personal distribution of these strains, they observed that the distribution of *P. acnes* strains was more individual-specific than site-specific, whereas *S. epidermidis* strains were more site-specific than individual-specific. Future investigations will need to focus on how the distribution of these strains varies over time and with changes in health.

The strength of metagenomic sequencing is the ability to survey the functional potential of microbial communities. To investigate this, Oh and colleagues compared their genomic data from each body site with reference genomes, which contain functional annotation for specific genes. Perhaps the most interesting result of this analysis was the identification of antibiotic-resistance genes that were specific to individuals and body sites. Appreciating the diversity and distribution of such genes across the skin could prove crucial in customizing therapies for the treatment of skin infections. More broadly, the authors were able to identify a strong functional signature between individuals, but found that its composition varied across the body. This result confirms the finding, from taxonomic analyses, that each body site provides a unique niche.

However, the limitation of metagenomic sequencing is that it describes only the functional potential of a community. As the researchers note, transcriptome analysis of the skin microbiota — defining the genes actually transcribed by the microorganisms — will be needed to identify the functional groups that are expressed at each site. It will be interesting to see whether populations such as *P. acnes*, which are found across the body, vary in their gene expression across the range of niches.



**Figure 1 | Skin partners.** Healthy human skin (cells shown in yellow) is colonized by a diverse array of microorganisms, including bacteria (magenta) and fungi (cyan). Scale bar, 10  $\mu$ m.

A frustrating but also exciting result of this analysis was the realization that between 2% and 96% of the sequence reads in each sample did not map to any of the reference genomes. Furthermore, many of the reads that did map could not be assigned a function on the basis of known genes. These results only underscore the individuality of the skin microbiota and beg for further cultivation and genome sequencing of skin-associated microbial populations. As comprehensive as this study was, the results demonstrate the need for a ‘multi-omic’ approach and time-series data. Sampling an individual over time would allow us to see how their particular microbiome varies in its composition and gene expression during transitions between health and disease. As this study indicates, cross-sectional studies are challenged by the enormous heterogeneity in the composition of the skin microbiota between individuals. Changes observed during such health–disease transitions might provide a better understanding of the relevance of these unknown sequences, which the researchers refer to as metagenomic dark matter. It is probable that this dark matter contains genes crucial to the functions that are unique to each niche and individual.

According to the ‘hygiene hypothesis’, our modern, sanitized world has fostered the spread of autoimmune disorders such

as allergies and asthma, by decreasing exposure to microorganisms during early life and thereby impeding the normal development of the immune system<sup>6</sup>. Just as probiotics and fibre (as a prebiotic) have emerged as consumer products designed to promote gut bacterial populations that are associated with health, it is tempting to interpret the data from Oh and colleagues as a call to develop similar products. For example, the presence of lipophilic *Corynebacterium* and *Malassezia* populations in the healthy people in this study suggests that moisturizing creams could be acting as a prebiotic to feed these organisms. With such knowledge, instead of reaching for a hand sanitizer that kills such populations, we might soon be able to reach for a product that fertilizes our skin microbiota to improve its ability to resist the colonization by potentially pathogenic organisms. ■

**Patrick D. Schloss** is in the Department of Microbiology & Immunology, University of Michigan, Ann Arbor, Michigan 48109-5620, USA.  
e-mail: pschloss@umich.edu

1. Oh, J. *et al.* *Nature* **514**, 59–64 (2014).
2. Grice, E. A. *et al.* *Science* **324**, 1190–1192 (2009).
3. Findley, K. *et al.* *Nature* **498**, 367–370 (2013).
4. Oh, J. *et al.* *Genome Res.* **23**, 2103–2114 (2013).
5. Oh, J., Conlan, S., Polley, E. C., Segre, J. A. & Kong, H. H. *Genome Med.* **4**, 77 (2012).
6. Strachan, D. P. *Br. Med. J.* **299**, 1259–1260 (1989).

# Reversion of advanced Ebola virus disease in nonhuman primates with ZMapp

Xiangguo Qiu<sup>1</sup>, Gary Wong<sup>1,2</sup>, Jonathan Audet<sup>1,2</sup>, Alexander Bello<sup>1,2</sup>, Lisa Fernando<sup>1</sup>, Judie B. Alimonti<sup>1</sup>, Hugues Fausther-Bovendo<sup>1,2</sup>, Haiyan Wei<sup>1,3</sup>, Jenna Aviles<sup>1</sup>, Ernie Hiatt<sup>4</sup>, Ashley Johnson<sup>4</sup>, Josh Morton<sup>4</sup>, Kelsi Swope<sup>4</sup>, Ognian Bohorov<sup>5</sup>, Natasha Bohorova<sup>5</sup>, Charles Goodman<sup>5</sup>, Do Kim<sup>5</sup>, Michael H. Pauly<sup>5</sup>, Jesus Velasco<sup>5</sup>, James Pettitt<sup>6†</sup>, Gene G. Olinger<sup>6†</sup>, Kevin Whaley<sup>5</sup>, Bianli Xu<sup>3</sup>, James E. Strong<sup>1,2,7</sup>, Larry Zeitlin<sup>5</sup> & Gary P. Kobinger<sup>1,2,8,9</sup>

**Without an approved vaccine or treatments, Ebola outbreak management has been limited to palliative care and barrier methods to prevent transmission. These approaches, however, have yet to end the 2014 outbreak of Ebola after its prolonged presence in West Africa. Here we show that a combination of monoclonal antibodies (ZMapp), optimized from two previous antibody cocktails, is able to rescue 100% of rhesus macaques when treatment is initiated up to 5 days post-challenge. High fever, viraemia and abnormalities in blood count and blood chemistry were evident in many animals before ZMapp intervention. Advanced disease, as indicated by elevated liver enzymes, mucosal haemorrhages and generalized petechia could be reversed, leading to full recovery. ELISA and neutralizing antibody assays indicate that ZMapp is cross-reactive with the Guinean variant of Ebola. ZMapp exceeds the efficacy of any other therapeutics described so far, and results warrant further development of this cocktail for clinical use.**

Ebola virus (EBOV) infections cause severe illness in humans, and after an incubation period of 3 to 21 days, patients initially present with general flu-like symptoms before a rapid progression to advanced disease characterized by haemorrhage, multiple organ failure and a shock-like syndrome<sup>1</sup>. In the spring of 2014, a new EBOV variant emerged in the West African country of Guinea<sup>2</sup>, an area in which EBOV had not been previously reported. Despite a sustained international response from local and international authorities including the Ministry of Health (MOH), World Health Organization (WHO) and Médecins Sans Frontières (MSF) since March 2014, the outbreak has yet to be brought to an end after five months. As of 15 August 2014, there are 2,127 total cases and 1,145 deaths spanning Guinea, Sierra Leone, Liberia and Nigeria<sup>3</sup>. So far, this outbreak has set the record for the largest number of cases and fatalities, in addition to geographical spread<sup>4</sup>. Controlling an EBOV outbreak of this magnitude has proven to be a challenge and the outbreak is predicted to last for at least several more months<sup>5</sup>. In the absence of licensed vaccines and therapeutics against EBOV, there is little that can be done for infected patients outside of supportive care, which includes fluid replenishment, administration of antivirals, and management of secondary symptoms<sup>6,7</sup>. With overburdened personnel, and strained local and international resources, experimental treatment options cannot be considered for compassionate use in an orderly fashion at the moment. However, moving promising strategies forward through the regulatory process of clinical development has never been more urgent.

Over the past decade, several experimental strategies have shown promise in treating EBOV-challenged nonhuman primates (NHPs) after infection. These include recombinant human activated protein C (rhAPC)<sup>8</sup>, recombinant nematode anticoagulant protein c2 (rNAPc2)<sup>9</sup>, small interfering RNA (siRNA)<sup>10</sup>, positively-charged phosphorodiamidate morpholino oligomers (PMOplus)<sup>11</sup>, the vesicular stomatitis virus vaccine (VSVΔG-EBOVGp)<sup>12</sup>, as well as the monoclonal antibody (mAb) cocktails

MB-003 (consisting of human or human–mouse chimaeric mAbs c13C6, h13F6 and c6D8)<sup>13</sup> and ZMAb (consisting of murine mAbs m1H3, m2G4 and m4G7)<sup>14</sup>. Of these, only the antibody-based candidates have demonstrated substantial benefits in NHPs when administered greater than 24 h past EBOV exposure. Follow-up studies have shown that MB-003 is partially efficacious when administered therapeutically after the detection of two disease ‘triggers’<sup>15</sup>, and ZMAb combined with an adenovirus-based adjuvant provides full protection in rhesus macaques when given up to 72 h after infection<sup>16</sup>.

The current objective is to develop a therapeutic superior to both MB-003 and ZMAb, which could be used for outbreak patients, primary health-care providers, as well as high-containment laboratory workers in the future. This study aims to first identify an optimized antibody combination derived from MB-003 and ZMAb components, before determining the therapeutic limit of this mAb cocktail in a subsequent experiment. To extend the antibody half-life in humans and to facilitate clinical acceptance, the individual murine antibodies in ZMAb were first chimeraized with human constant regions (cZMAb; components: c1H3, c2G4 and c4G7). The cZMAb components were then produced in *Nicotiana benthamiana*<sup>17</sup>, using the large-scale, Current Good Manufacturing Practice-compatible Rapid Antibody Manufacturing Platform (RAMP) and magnICON vectors that currently also manufactures the individual components of cocktail MB-003, before efficacy testing in animals.

## Selecting for the best mAb combinations

Our efforts to down-select for an improved mAb cocktail comprising components of MB-003 and ZMAb began with the testing of individual MB-003 antibodies in guinea pigs and NHPs. In guinea pig studies, animals were given one dose of mAb c13C6, h13F6 or c6D8 individually (totaling 5 mg per animal) at 1 day post-infection (dpi) with 1,000 × LD<sub>50</sub> (median lethal dose) of guinea-pig-adapted EBOV, Mayinga variant

<sup>1</sup>National Laboratory for Zoonotic Diseases and Special Pathogens, Public Health Agency of Canada, Winnipeg, Manitoba R3E 3R2, Canada. <sup>2</sup>Department of Medical Microbiology, University of Manitoba, Winnipeg, Manitoba R3E 0J9, Canada. <sup>3</sup>Institute of Infectious Disease, Henan Centre for Disease Control and Prevention, Zhengzhou, 450012 Henan, China. <sup>4</sup>Kentucky BioProcessing, Owensboro, Kentucky 42301, USA. <sup>5</sup>Mapp Biopharmaceutical Inc., San Diego, California 92121, USA. <sup>6</sup>United States Army Medical Research Institute of Infectious Diseases (USAMRIID), Frederick, Maryland 21702, USA. <sup>7</sup>Department of Pediatrics and Child Health, University of Manitoba, Winnipeg, Manitoba R3A 1S1, Canada. <sup>8</sup>Department of Immunology, University of Manitoba, Winnipeg, Manitoba R3E 0T5, Canada. <sup>9</sup>Department of Pathology and Laboratory Medicine, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA. <sup>†</sup>Present address: Integrated Research Facility, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Frederick, Maryland 21702, USA.



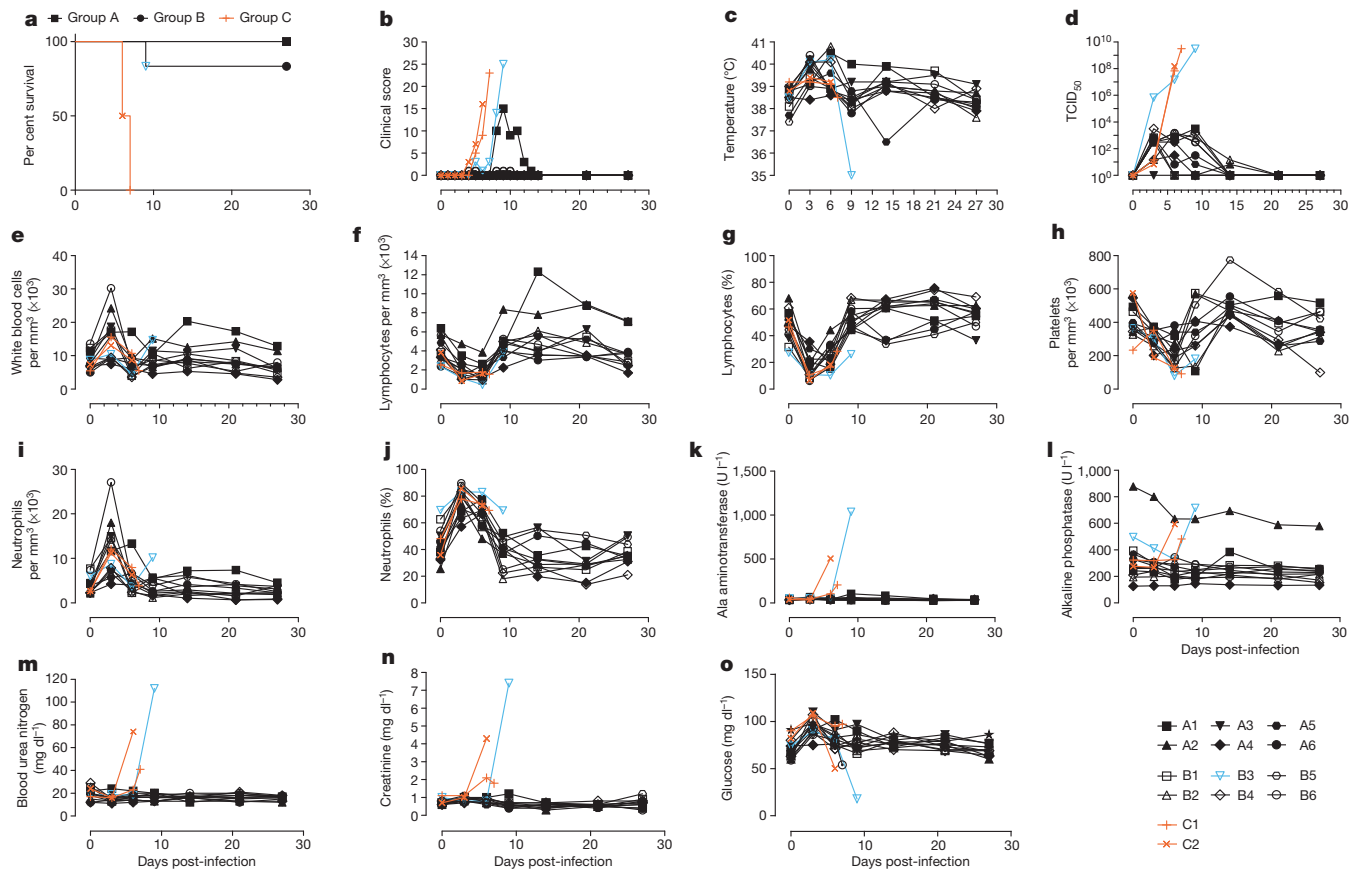
**Table 1 | Efficacy of individual and combined monoclonal antibody treatments in guinea pigs and nonhuman primates**

Treatment groups, time of treatment	Dose (mg)	Mean time to death (days $\pm$ s.d.)	No. survivors/total	Survival (%)	Weight loss (%)	<i>P</i> value, compared with	
						cZMAb	MB-003
<b>Guinea pigs</b>							
PBS, 3 dpi	N/A	7.3 $\pm$ 0.5	0/4	0	9%	—	—
cZMAb, 3 dpi	5	11.6 $\pm$ 1.8	1/6	17	7%	—	—
MB-003, 3 dpi	5	8.2 $\pm$ 1.5	0/6	0	40%	—	—
ZMapp1, 3 dpi	5	9.0 $\pm$ 0.0	4/6	67	<5%	0.190	0.0147
ZMapp2, 3 dpi	5	8.3 $\pm$ 0.6	3/6	50	8%	0.634	0.0692
ZMapp3, 3 dpi	5	8.6 $\pm$ 1.1	1/6	17	9%	0.224	0.411
c13C6, 1 dpi	5	8.4 $\pm$ 1.7	1/6	17	9%	—	—
h13F6, 1 dpi	5	10.2 $\pm$ 1.8	1/6	17	21%	—	—
c6D8, 1 dpi	5	10.5 $\pm$ 2.2	0/6	0	38%	—	—
<b>Nonhuman primates</b>							
PBS, 1 dpi	N/A	8.4 $\pm$ 1.9	0/1	0			
MB-003, 1 dpi	50	14.0 $\pm$ 2.8	1/3	33			
c13C6, 1 dpi	50	9.0 $\pm$ 1.4	1/3	33			
h13F6, 1 dpi	50	9.0 $\pm$ 2.0	0/3	0			
c6D8, 1 dpi	50	9.7 $\pm$ 0.6	0/3	0			

(EBOV-M-GPA). Survival and weight loss were monitored over 28 days. Treatment with c13C6 or h13F6 yielded 17% survival (1 of 6 animals) with a mean time to death of  $8.4 \pm 1.7$  and  $10.2 \pm 1.8$  days, respectively. The average weight loss for c13C6 or h13F6-treated animals was 9% and 21% (Table 1). In nonhuman primates, animals were given three doses of mAb c13C6, h13F6 or c6D8, beginning at 24 h after challenge with the Kikwit variant of EBOV (EBOV-K)<sup>18</sup>, and survival was monitored over 28 days. Only c13C6 treatment yielded any survivors, with 1 of 3 animals protected from EBOV challenge (Table 1), confirming in two

separate animal models that c13C6 is the component that provides the highest level of protection in the MB-003 cocktail.

We then tested mAb c13C6 in combination with two of three mAbs from ZMAb in guinea pigs. The individual antibodies composing ZMAb were originally chosen for protection studies based on their *in vivo* protection of guinea pigs against EBOV-M-GPA<sup>19</sup>, and all three possible combinations were tested: ZMapp1 (c13C6+c2G4+c4G7), ZMapp2 (c13C6+c1H3+c2G4) and ZMapp3 (c13C6+c1H3+c4G7), and compared to the originator cocktails ZMAb and MB-003. Three days after



**Figure 1 | Post-exposure protection of EBOV-infected nonhuman primates with ZMapp1 and ZMapp2.** Rhesus macaques were challenged with EBOV-K, and 50 mg kg<sup>-1</sup> of ZMapp1 (Group A) or ZMapp2 (Group B) were administered on days 3, 6, and 9 ( $n = 6$  per treatment group,  $n = 2$  for controls). Non-specific IgG mAb or PBS was administered as a control (Group C). **a**, Kaplan-Meier survival curves (log-rank tests: Group A vs Group C

$P = 0.0039$ ; Group B vs Group C  $P = 0.0039$ ). **b**, Clinical score. **c**, Rectal temperature. **d**, EBOV viraemia by TCID<sub>50</sub>. Blood parameters: **e**, white blood cell count; **f**, lymphocyte count; **g**, lymphocyte percentage; **h**, platelet count; **i**, neutrophil count; **j**, neutrophil percentage; **k**, alanine aminotransferase; **l**, alkaline phosphatase; **m**, blood urea nitrogen; **n**, creatinine; **o**, glucose.

challenge with  $1,000 \times \text{LD}_{50}$  of EBOV-M-GPA, the animals received a single combined dose of 5 mg of antibodies. This dosage is purposely given to elicit a suboptimal level of protection with the cZMAb and MB-003 cocktails, such that potential improvements with the optimized mAb combinations can be identified. Of the tested cocktails, ZMapp1 showed the best protection, with 4 of 6 survivors and less than 5% average weight loss (Table 1). ZMapp2 was next with 3 of 6 survivors and 8% average weight loss, and ZMapp3 protected 1 of 6 animals (Table 1). The level of protection afforded by ZMapp3 was not a statistically significant increase over cZMAb ( $P = 0.224$ , log-rank test compared to ZMAb,  $\chi^2 = 1.479$ , degrees of freedom (d.f.) = 1), and showed the same survival rate along with a similar average weight loss (Table 1). As a result, only ZMapp1 and ZMapp2 were carried forward to NHP studies.

### ZMapp1 or ZMapp2-treated NHPs

Rhesus macaques were used to determine whether administration of ZMapp1 or ZMapp2 was superior to ZMAb and MB-003 in terms of extending the treatment window. Owing to mAb availability constraints, m4G7 was used in place of c4G7 for this NHP experiment. The experiment consisted of six NHPs per group receiving three doses of ZMapp1 (Group A) or ZMapp2 (Group B) at  $50 \text{ mg kg}^{-1}$  intravenously at 3-day intervals, beginning 3 days after a lethal intramuscular challenge with  $4,000 \times$  median tissue culture infective dose ( $\text{TCID}_{50}$ ) (or 2,512 plaque-forming units (p.f.u.)) of EBOV-K. Control animals were given phosphate-buffered saline (PBS) or mAb 4E10 (C1 and C2, respectively). Mock-treated animals succumbed to disease between 6–7 dpi with symptoms typical

of EBOV (Fig. 1a), characterized by high clinical scores but no fever (Fig. 1b, c), in addition to viral titres up to approximately  $10^8$  and  $10^9$   $\text{TCID}_{50}$  by the time of death (Fig. 1d). Analysis of blood counts and serum biochemistry revealed leukocytopenia, thrombocytopenia, severe rash, decreased levels of glucose, as well as increased levels of alkaline phosphatase, alanine aminotransferase, blood urea nitrogen and creatinine at end-stage EBOV disease (Fig. 1e–o, Table 2).

All six Group A NHPs survived the challenge with mild signs of disease (Fig. 1a, Table 2) ( $P = 0.0039$ , log-rank test,  $\chi^2 = 8.333$ , d.f. = 1, comparing to Group C), with the exception of A1 which showed an elevated clinical score (Fig. 1b), increased levels of alanine aminotransferase, total bilirubin, and decreased phosphate (Fig. 1, Table 2). However, this animal recovered after the third ZMapp1 dose and the clinical score dropped to zero by 15 dpi (Fig. 1b). A fever was detected in all but one of the NHPs (A4) at 3 dpi, the start of the first ZMapp1 dose (Fig. 1c). Viraemia was also detected beginning at 3 dpi by  $\text{TCID}_{50}$  in all but one animal from blood sampled just before the administration of treatment (A3) (Fig. 1d), and similar results were observed by quantitative PCR with reverse transcription (RT-qPCR, Extended Data Table 1). The viraemia decreased to undetectable levels by 21 dpi. EBOV shedding was not detected from oral, nasal and rectal swabs by RT-qPCR in any of the Group A animals (Extended Data Tables 2–4).

For Group B, 5 of 6 NHPs survived with B3 succumbing to disease at 9 dpi (Fig. 1a) ( $P = 0.0039$ , log-rank test,  $\chi^2 = 8.333$ , d.f. = 1, comparing to Group C). Surviving animals showed only mild signs of disease (Table 2). The moribund animal showed increased clinical scores (Fig. 1b),

**Table 2 | Clinical findings of EBOV-infected NHPs from 1 to 27 dpi**

Animal ID	Treatment group	Clinical findings					Outcome
		Body temperature	Rash	White blood cells	Platelets	Biochemistry	
<b>A1</b>	$50 \text{ mg kg}^{-1}$ c13C6+c2G4+m4G7, 3 dpi	Fever (6, 9, 14 dpi)			Thrombocytopenia (6, 9 dpi)	ALT↑ (9, 14 dpi), TBIL↑ (9 dpi), PHOS↓ (6 dpi)	Survived
<b>A2</b>	$50 \text{ mg kg}^{-1}$ c13C6+c2G4+m4G7, 3 dpi	Fever (3 dpi)		Leukocytosis (3 dpi)		CRE↓ (14 dpi)	Survived
<b>A3</b>	$50 \text{ mg kg}^{-1}$ c13C6+c2G4+m4G7, 3 dpi	Fever (3 dpi)		Leukocytosis (3 dpi)	Thrombocytopenia (6 dpi)		Survived
<b>A4</b>	$50 \text{ mg kg}^{-1}$ c13C6+c2G4+m4G7, 3 dpi			Leukocytopenia (9 dpi)	Thrombocytopenia (3, 6, 14, 21, 27 dpi)		Survived
<b>A5</b>	$50 \text{ mg kg}^{-1}$ c13C6+c2G4+m4G7, 3 dpi	Fever (3, 6, 9 dpi)		Leukocytopenia (9 dpi)	Thrombocytopenia (3, 21 dpi)		Survived
<b>A6</b>	$50 \text{ mg kg}^{-1}$ c13C6+c2G4+m4G7, 3 dpi	Fever (3 dpi)					Survived
<b>B1</b>	$50 \text{ mg kg}^{-1}$ ZMapp2, 3 dpi	Fever (3, 14, 21 dpi)		Leukocytopenia (6, 14, 21, 27 dpi)	Thrombocytopenia (6 dpi)		Survived
<b>B2</b>	$50 \text{ mg kg}^{-1}$ ZMapp2, 3 dpi	Fever (3, 6 dpi)			Thrombocytopenia (6, 9 dpi)		Survived
<b>B3</b>	$50 \text{ mg kg}^{-1}$ ZMapp2, 3 dpi	Fever (3, 6 dpi), Hypothermia (9 dpi)	Severe rash (9 dpi)		Thrombocytopenia (6, 9 dpi)	ALT↑↑↑ (9 dpi), TBIL↑↑ (9 dpi), BUN↑↑↑ (9 dpi), CRE↑↑↑ (9 dpi), GLU↓ (9 dpi)	Died, 9 dpi
<b>B4</b>	$50 \text{ mg kg}^{-1}$ ZMapp2, 3 dpi	Fever (3, 6 dpi)		Leukocytopenia (6 dpi)	Thrombocytopenia (6, 27 dpi)		Survived
<b>B5</b>	$50 \text{ mg kg}^{-1}$ ZMapp2, 3 dpi	Fever (3, 6, 14, 21 dpi)		Leukocytosis (3 dpi)	Thrombocytopenia (3, 6 dpi)		Survived
<b>B6</b>	$50 \text{ mg kg}^{-1}$ ZMapp2, 3 dpi	Fever (3 dpi)		Leukocytosis (3 dpi), Leukocytopenia (6, 9, 14, 21, 27 dpi)	Thrombocytopenia (6 dpi)	PHOS↓ (3 dpi), CRE↓ (27 dpi)	Survived
<b>C1</b>	PBS, 3 dpi		Moderate rash (6 dpi), Severe rash (7 dpi)	Leukocytosis (3 dpi)	Thrombocytopenia (6, 7 dpi)	ALB↓ (7 dpi), ALT↑ (7 dpi), BUN↑ (7 dpi)	Died, 7 dpi
<b>C2</b>	Control mAb, 3 dpi		Severe rash (6 dpi)	Leukocytopenia (6, 7 dpi)	Thrombocytopenia (6, 7 dpi)	ALP↑ (3 dpi), ALT↑↑↑ (6 dpi), BUN↑ (6 dpi), CRE↑↑↑ (6 dpi)	Died, 6 dpi

Hypothermia was defined as below  $35^\circ\text{C}$ . Fever was defined as  $>1.0^\circ\text{C}$  higher than baseline. Mild rash was defined as focal areas of petechiae covering  $<10\%$  of the skin, moderate rash as areas of petechiae covering 10 to 40% of the skin, and severe rash as areas of petechiae and/or ecchymosis covering  $>40\%$  of the skin. Leukocytopenia and thrombocytopenia were defined as a  $>30\%$  decrease in numbers of white blood cells and platelets, respectively. Leukocytosis and thrombocytosis were defined as a twofold or greater increase in numbers of white blood cells and platelets over baseline, where white blood cell count  $> 11 \times 10^3$ . ↑, two- to threefold increase; ↑↑, four- to fivefold increase; ↑↑↑, greater than fivefold increase; ↓, two- to threefold decrease; ↓↓, four- to fivefold decrease; ↓↓↓, greater than fivefold decrease. ALB, albumin; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AMY, amylase; TBIL, total bilirubin; BUN, blood urea nitrogen; PHOS, phosphate; CRE, creatinine; GLU, glucose; GLOB, globulin.



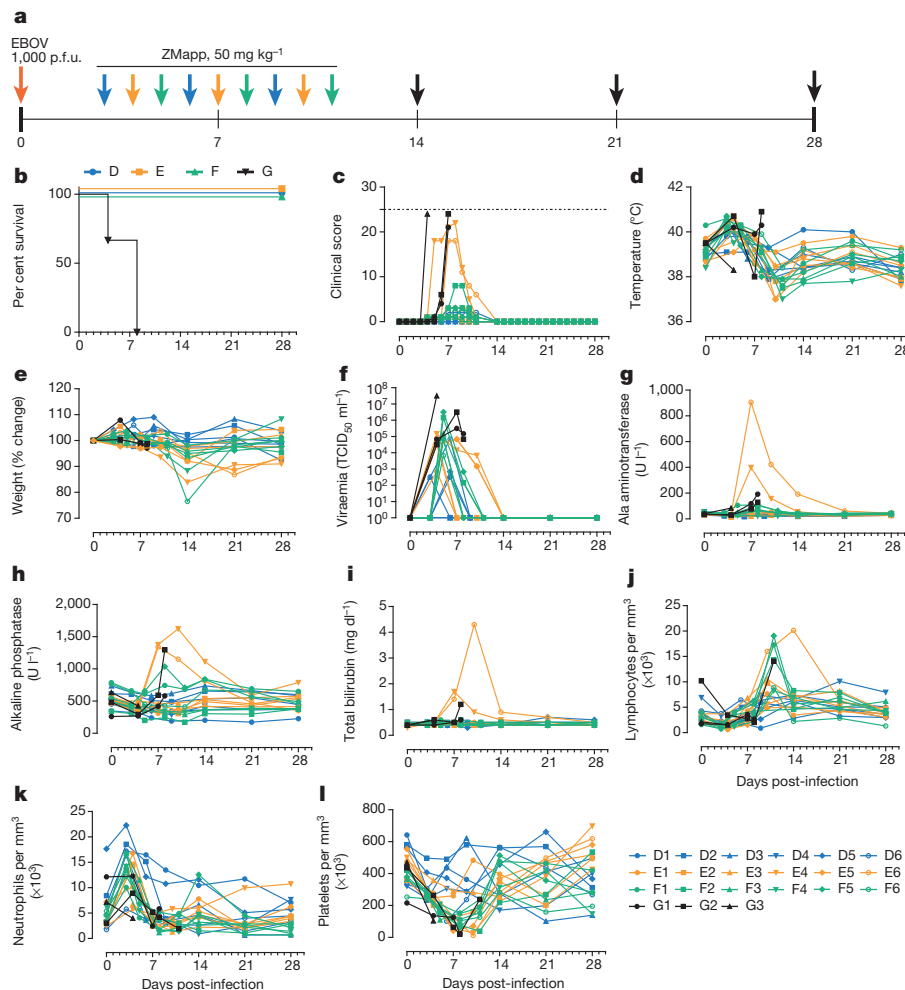
in addition to a drastic drop in body temperature shortly before death (Fig. 1c). At the time of death, animal B3 had elevated levels of alanine aminotransferase, total bilirubin, blood urea nitrogen and creatinine, in addition to decreased levels of glucose, suggesting multiple organ failure (Fig. 1). All six Group B animals showed fever in addition to viraemia at 3 dpi by TCID<sub>50</sub> and RT-qPCR (Fig. 1d, Extended Data Table 1). It was interesting to note that in B3, the viraemia reached approximately 10<sup>6</sup> TCID<sub>50</sub> after 3 dpi (Fig. 1d), suggesting that this NHP was particularly susceptible to EBOV infection. No escape mutants were detected with this animal. The administration of ZMapp2 at the reported concentrations was unable to effectively control viraemia at this level. Virus shedding was also detected from the oral and rectal swabs by RT-qPCR in the moribund NHP B3 (Extended Data Tables 2–4). Since ZMapp1 demonstrated superior protection to ZMapp2 in this survival study, ZMapp1 (now trademarked as ZMapp by MappBio Pharmaceuticals) was carried forward to test the limits of protection conferred by this mAb cocktail in a subsequent investigation.

### ZMapp-treated NHPs

In this experiment, rhesus macaques were assigned into three treatment groups of six and a control group of three animals, with all treatment NHPs receiving three doses of ZMapp (c13C6 + c2G4 + c4G7, 50 mg kg<sup>-1</sup> per dose) spaced 3 days apart. After a lethal intramuscular challenge with

1,000 × TCID<sub>50</sub> (or 628 p.f.u.) of EBOV-K<sup>18</sup>, we treated the animals with ZMapp at 3, 6 and 9 dpi (Group D); 4, 7, and 10 dpi (Group E); or 5, 8 and 11 dpi (Group F). The control animals (Group G) were given mAb 4E10 as an IgG isotype control (*n* = 1) or PBS (*n* = 2) in place of ZMapp starting at 4 dpi (Fig. 2a). All animals treated with ZMapp survived the infection, whereas the three control NHPs (G1, G2 and G3) succumbed to EBOV-K infection at 4, 8 and 8 dpi, respectively (*P* = 3.58 × 10<sup>-5</sup>, log-rank test,  $\chi^2$  = 23.25, d.f. = 3, comparing all groups) (Fig. 2b). At the time ZMapp treatment was initiated, fever, leukocytosis, thrombocytopenia and viraemia could be detected in the majority of the animals (Fig. 2c–f, Table 3, Extended Data Table 5). All animals presented with detectable abnormalities in blood counts and serum biochemistry during the course of the experiment (Fig. 2g–l, Table 3).

The Group F animals did not seem to be as sick as animals E4 and E6 on the basis of clinical scores (Fig. 2c, Extended Data Fig. 1), both animals E4 and E6 were near the clinical limit for IACUC mandated euthanasia at 5 and 7 dpi, respectively. Animal E4 had a flushed face and severe rash on more than 40% of its body surface between 5 and 8 dpi in addition to nasal haemorrhage at 7 dpi, and animal E6 had a flushed face and petechiae on its arms and legs between 7 and 9 dpi, in addition to jaundice between 10 and 14 dpi. This indicates that host genetic factors may have a role in the differential susceptibility of individual NHPs to EBOV-K infections. Fever, leukocytosis, thrombocytopenia and a



**Figure 2 | Post-exposure protection of EBOV-infected nonhuman primates with ZMapp.** a–f, Rhesus macaques (*n* = 6 per ZMapp treatment group, *n* = 3 for controls) were challenged with EBOV-K, and 50 mg kg<sup>-1</sup> of ZMapp were administered beginning at 3 (Group D), 4 (Group E) or 5 (Group F) days after challenge. Non-specific IgG mAb or PBS was administered as a control (Group G). a, Timeline of infection, treatment and sample days. b, Kaplan–Meier survival curves (log-rank test: overall comparison *P* = 3.58 × 10<sup>-5</sup>).

c, Clinical scores; the dashed line indicates the minimum score requiring mandatory euthanasia. d, Rectal temperature. e, Percentage body weight change. f, EBOV viraemia by TCID<sub>50</sub>. g–l, Selected clinical parameters of Group A to D animals. g, Alanine aminotransferase; h, alkaline phosphatase; i, total bilirubin. j–l, Counts for lymphocytes (j), neutrophils (k) and platelets (l) over the course of the experiment.

**Table 3 | Clinical findings of EBOV-infected NHPs from 1 to 28 dpi**

Animal ID	Treatment group	Clinical findings Body temperature	Rash	White blood cells	Platelets	Biochemistry	Outcome
<b>D1</b>	50 mg kg <sup>-1</sup> ZMapp, 3 dpi	Fever (3, 6, 14, 21 dpi)		Leukocytosis (3, 6, 21 dpi)	Thrombocytopenia (3, 6, 9, 14, 21 dpi)	ALB↓ (14, 21 dpi), ALP↓ (9, 14, 21, 28 dpi), AMY↓ (9 dpi), GLOB↑ (21, 28 dpi), PHOS↓ (9 dpi)	Survived
<b>D2</b>	50 mg kg <sup>-1</sup> ZMapp, 3 dpi			Leukocytopenia (21, 28 dpi)	Thrombocytopenia (28 dpi)		Survived
<b>D3</b>	50 mg kg <sup>-1</sup> ZMapp, 3 dpi	Fever (3 dpi)		Leukocytosis (3, 14 dpi)	Thrombocytopenia (3, 21, 28 dpi)	ALT↓ (6 dpi)	Survived
<b>D4</b>	50 mg kg <sup>-1</sup> ZMapp, 3 dpi			Leukocytopenia (14 dpi)	Thrombocytopenia (14, 21 dpi)	ALT↓ (9 dpi), CRE↑ (14 dpi)	Survived
<b>D5</b>	50 mg kg <sup>-1</sup> ZMapp, 3 dpi	Fever (3 dpi)		Leukocytopenia (21, 28 dpi)	Thrombocytopenia (6, 9 dpi)	ALB↓ (9 dpi), BUN↓ (3, 6, 14, 21, 28 dpi)	Survived
<b>D6</b>	50 mg kg <sup>-1</sup> ZMapp, 3 dpi				Thrombocytopenia (6 dpi)		Survived
<b>E1</b>	50 mg kg <sup>-1</sup> ZMapp, 4 dpi				Thrombocytopenia (4, 7, 21 dpi)	AMY↓ (4, 21 dpi), AMY↓ (7, 10, 14 dpi), CRE↓ (21, 28 dpi)	Survived
<b>E2</b>	50 mg kg <sup>-1</sup> ZMapp, 4 dpi	Fever (4 dpi)		Leukocytosis (4, 10 dpi)	Thrombocytopenia (4, 7, 10, 21 dpi)	ALT ↓↓ (4 dpi), GLU↑ (4 dpi)	Survived
<b>E3</b>	50 mg kg <sup>-1</sup> ZMapp, 4 dpi	Fever (4 dpi)		Leukocytosis (4, 10 dpi)	Thrombocytopenia (7, 10, 14 dpi)	CRE↓ (14 dpi)	Survived
<b>E4</b>	50 mg kg <sup>-1</sup> ZMapp, 4 dpi		Severe rash (5, 6, 7, 8 dpi), Mild rash (9 dpi)	Leukocytosis (10, 14, 21, 28 dpi)	Thrombocytopenia (4, 7, 10, 14 dpi)	ALP↑ (7, 10, 14 dpi), ALT ↑↑↑ (7 dpi), ALT ↑↑ (10 dpi), AMY↓ (4, 7, 10 dpi), TBIL↑↑↑ (7 dpi), TBIL↑ (10, 14 dpi), PHOS↓ (7, 10 dpi), K <sup>+</sup> ↓ (4 dpi)	Survived
<b>E5</b>	50 mg kg <sup>-1</sup> ZMapp, 4 dpi	Fever (7 dpi)		Leukocytosis (4 dpi)	Thrombocytopenia (4, 7, 10, 14 dpi)	ALT↑ (7 dpi), AMY↓ (4, 7 dpi), PHOS↓ (10 dpi)	Survived
<b>E6</b>	50 mg kg <sup>-1</sup> ZMapp, 4 dpi	Fever (4 dpi)	Mild rash (7, 8, 9 dpi)	Leukocytosis (4, 10, 14 dpi)	Thrombocytopenia (4, 7, 10, 14 dpi)	ALP↑ (7, 10 dpi), ALT ↑↑↑ (7, 10, 14 dpi), AMY↓ (7, 10 dpi), TBIL↑↑ (7 dpi), TBIL↑↑↑ (10 dpi), TBIL↑ (14 dpi), PHOS↓ (7 dpi), GLOB↑ (21 dpi)	Survived
<b>F1</b>	50 mg kg <sup>-1</sup> ZMapp, 5 dpi			Leukocytosis (11 dpi)	Thrombocytopenia (3, 5, 8, 11 dpi)	AMY↓ (5 dpi), PHOS↓ (11 dpi), CRE↓ (28 dpi)	Survived
<b>F2</b>	50 mg kg <sup>-1</sup> ZMapp, 5 dpi	Fever (3, 5 dpi)	Mild rash (8 dpi)	Leukocytosis (3, 5, 11 dpi)	Thrombocytopenia (3, 5, 8, 11, 14, 21 dpi)	PHOS↓ (11 dpi), CRE↓ (11 dpi)	Survived
<b>F3</b>	50 mg kg <sup>-1</sup> ZMapp, 5 dpi			Leukocytopenia (8 dpi), Leukocytosis (3 dpi)	Thrombocytopenia (5, 8, 11, 21 dpi)	ALT↑ (8 dpi), CRE↓ (14 dpi)	Survived
<b>F4</b>	50 mg kg <sup>-1</sup> ZMapp, 5 dpi	Fever (3, 5 dpi)		Leukocytopenia (8 dpi)	Thrombocytopenia (5, 8, 11, 28 dpi)	PHOS↓ (8 dpi)	Survived
<b>F5</b>	50 mg kg <sup>-1</sup> ZMapp, 5 dpi	Fever (3 dpi)		Leukocytosis (3, 11, 14 dpi)	Thrombocytopenia (5, 8, 11 dpi)	PHOS↓ (5, 8 dpi), CRE↓ (8, 11, 21, 28 dpi)	Survived
<b>F6</b>	50 mg kg <sup>-1</sup> ZMapp, 5 dpi	Fever (3 dpi)		Leukocytopenia (8, 21, 28 dpi)	Thrombocytopenia (8, 11, 21 dpi)	PHOS↓ (5, 8, 11 dpi), GLU↑ (5 dpi)	Survived
<b>G1</b>	PBS, 4 dpi		Severe rash (4 dpi)	Leukocytopenia (4 dpi)	Thrombocytopenia (4 dpi)	AMY↓ (4 dpi)	Died, 4 dpi
<b>G2</b>	Control mAb, 4 dpi		Severe rash (8 dpi)	Leukocytopenia (7, 8 dpi)	Thrombocytopenia (4, 7, 8 dpi)	ALP↑ (8 dpi), ALT↑ (7 dpi), ALT ↑↑↑ (8 dpi), CRE↑ (8 dpi)	Died, 8 dpi
<b>G3</b>	PBS, 4 dpi	Fever (4, 8 dpi)	Severe rash (8 dpi)	Leukocytopenia (7, 8 dpi)	Thrombocytopenia (4, 7, 8 dpi)	ALP↑ (8 dpi), ALT↑ (7, 8 dpi), AMY↓ (7 dpi), AMY ↓↓ (8 dpi), TBIL↑ (8 dpi), PHOS↓ (7 dpi)	Died, 8 dpi

Hypothermia was defined as below 35 °C. Fever was defined as >1.0 °C higher than baseline. Mild rash was defined as focal areas of petechiae covering <10% of the skin, moderate rash was defined as areas of petechiae covering 10 to 40% of the skin, and severe rash was defined as areas of petechiae and/or ecchymosis covering >40% of the skin. Leukocytopenia and thrombocytopenia were defined as a >30% decrease in the numbers of white blood cells and platelets, respectively. Leukocytosis and thrombocytosis were defined as a twofold or greater increase in numbers of white blood cells and platelets above baseline, where white blood cell count > 11 × 10<sup>3</sup>. ↑, two- to threefold increase; ↑↑, four- to fivefold increase; ↑↑↑, greater than fivefold increase; ↓, two- to threefold decrease; ↓↓, four- to fivefold decrease; ↓↓↓, greater than fivefold decrease. ALB, albumin; ALP, alkaline phosphatase; ALT, alanine aminotransferase; AMY, amylase; TBIL, total bilirubin; BUN, blood urea nitrogen; PHOS, phosphate; CRE, creatinine; GLU, glucose; K<sup>+</sup>, potassium; GLOB, globulin.



severe rash symptomatic of EBOV disease progression were detected in both E4 and E6 (Table 3). Increases in the level of liver enzymes alanine aminotransferase (10- to 30-fold increase), alkaline phosphatase (two- to threefold), and total bilirubin (3- to 11-fold) indicate significant liver damage (Fig. 2g–i), a hallmark of filovirus infections. However, ZMapp was successful in reversing observed disease symptoms and physiological abnormalities after 12 dpi, 2 days after the last ZMapp administration (Table 3). Furthermore, ZMapp treatment was able to lower the high virus loads observed in animals F2 and F5 (up to  $10^6$  TCID<sub>50</sub> ml<sup>-1</sup>) to undetectable levels by 14 dpi (Fig. 2f, Extended Data Fig. 2).

### ZMapp cross-reacts with Guinea EBOV

Although the results were very promising with EBOV-K-infected NHPs, it was unknown whether therapy with ZMapp would be similarly effective against the Guinean variant of EBOV (EBOV-G), the virus responsible for the West African outbreak. Direct comparison of published amino acid sequences between EBOV-G and EBOV-K showed that the epitopes targeted by ZMapp<sup>20,21</sup> were not mutated between the two virus variants (Fig. 3a), indicating that the antibodies should retain their specificity for the viral glycoprotein. To confirm this, *in vitro* assays were carried out to compare the binding affinity of c13C6, c2G4 and c4G7 to sucrose-purified EBOV-G and EBOV-K. As measured by enzyme-linked immunosorbent assay (ELISA), the ZMapp components showed slightly better binding kinetics for EBOV-G than for EBOV-K (Fig. 3b). Additionally, the neutralizing activity of individual mAbs was evaluated in the absence of complement for c2G4 and c4G7, and in the presence of complement for c13C6, as they have previously been shown to neutralize EBOV only under these conditions<sup>13</sup> (Fig. 3c). The results supported the ELISA binding data, with comparable neutralizing activities between the two viruses.

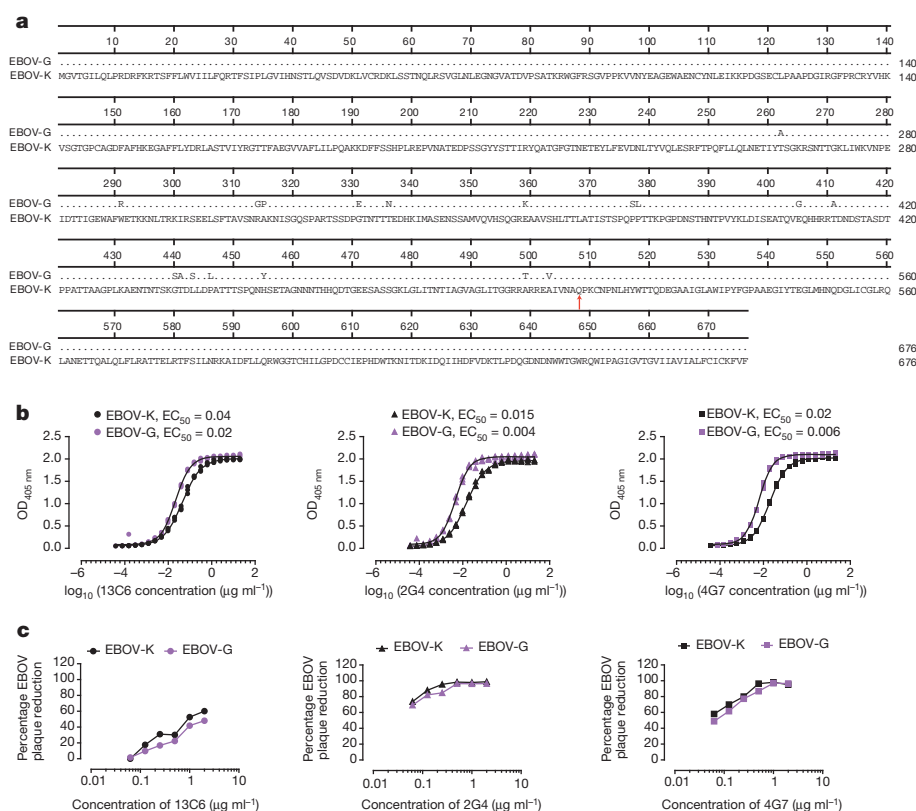
### Discussion

The West African outbreak of 2014 has highlighted the troubling absence of available vaccine or therapeutic options to save thousands of lives and stop the spread of EBOV. The lack of a clinically acceptable treatment offers limited incentive for people who suspect they might be infected to

report themselves for medical help. Several previous studies have showed that antibodies are crucial for host survival from EBOV<sup>22–24</sup>. Prior NHP studies have also demonstrated that the ZMAb cocktail could protect 100% or 50% of animals when dosing was initiated 1 or 2 dpi, while the MB-003 cocktail protected 67% of animals with the same dosing regimen. Before the success with monoclonal-antibody-based therapies, other candidate therapeutics had only demonstrated efficacy when given within 60 min of EBOV exposure.

Our results with ZMapp, a cocktail comprising of individual monoclonal antibodies selected from MB-003 and ZMAb, demonstrate for the first time the successful protection of NHPs from EBOV disease when intervention was initiated as late as 5 dpi. In the preceding ZMapp1/ZMapp2 experiment, 11 out of 12 treated animals had detectable fever (with the exception of A4), and live virus could be detected in the blood of 11 out of 12 animals (with the exception of A3) by 3 dpi. Therefore, for the majority of these animals, treatment was therapeutic (as opposed to post-exposure prophylaxis), initiated after two detectable triggers of disease. ZMapp2 was able to protect 5 of 6 animals when administered at 3 dpi. For reasons currently unknown, the lone non-survivor (B3) experienced a viraemia of  $10^6$  TCID<sub>50</sub> at 3 dpi, which is 100-fold greater than all other NHPs and approximately tenfold higher than that which ZMAb has been reported to suppress in a previous study<sup>16</sup>. This indicates enhanced EBOV replication in this animal, possibly owing to host factors. It is important to note that, despite the high levels of live circulating virus detected in B3, ZMapp2 administration was still able to prolong the life of this animal to 9 dpi, and suggests that in cases of high viraemia such as this, the dosage of monoclonal antibodies should be increased.

The highlight of these experimental results is undoubtedly ZMapp, which was able to reverse severe EBOV disease as indicated by the elevated liver enzymes, mucosal haemorrhages and rash in animals E4 and E6. The high viraemia (up to  $10^6$  TCID<sub>50</sub> ml<sup>-1</sup> of blood in some animals at the time of intervention) could also be effectively controlled without the presence of escape mutants, leading to full recovery of all treated NHPs by 28 dpi. In the absence of direct evidence demonstrating ZMapp efficacy against lethal EBOV-G infection in NHPs, results from ELISA



**Figure 3 | Amino acid alignment of the Kikwit and Guinea variants of EBOV, and *in vitro* antibody assays of mAbs c13C6, c2G4 and c4G7 with EBOV-G or EBOV-K virions.** **a**, Sequence alignment of the EBOV glycoprotein from the Kikwit (EBOV-K) and Guinea (EBOV-G) variants, with the binding epitopes of ZMapp indicated with an arrow. **b**, ELISA, note that for each antibody, the median effective concentrations (EC<sub>50</sub>) are different ( $P < 0.05$ , regression analysis) between the two antigens. **c**, Neutralizing antibody assay showing the activity of the individual mAbs composing ZMapp against EBOV-K (black) and EBOV-G (purple), the samples were run in triplicate.

and neutralizing antibody assays show that binding specificity is not abrogated between EBOV-K and EBOV-G, and therefore the levels of protection should not be affected. The compassionate use of ZMapp in two infected American healthcare workers, with apparently positive results pertaining to survival and reversion of EBOV disease<sup>25</sup>, may support this assertion. Rhesus macaques have approximately 55–80 ml of blood per kg of body weight<sup>26</sup>; at a dose of 50 mg kg<sup>-1</sup> of antibodies, the estimated starting concentration is approximately 625–909 µg ml<sup>-1</sup> of blood (total; ~200–300 µg ml<sup>-1</sup> for each antibody). Therefore, the low median effective concentration (EC<sub>50</sub>) values for EBOV-G (0.004–0.02 µg ml<sup>-1</sup>) bode well for treating EBOV-G infections with ZMapp.

Since the host antibody response is known to correlate with and is required for protection from EBOV infections<sup>23,24</sup>, monoclonal-antibody-based treatments are likely to form the centrepiece of any future therapeutic strategies for fighting EBOV outbreaks. However, whether ZMapp-treated survivors can be susceptible to re-infection is unknown. In a previous study of murine ZMAB-treated, EBOV-challenged NHP survivors, a re-challenge of these animals with the same virus at 10 and 13 weeks after initial challenge yielded 6 of 6 survivors and 4 of 6 survivors, respectively<sup>27</sup>. While specific CD4<sup>+</sup> and CD8<sup>+</sup> T-cell responses could be detected in all animals, the circulating levels of glycoprotein (GP)-specific IgG were shown to be tenfold lower in non-survivors compared to survivors, suggesting that antibody levels may be indicative of protective immunity<sup>27</sup>. Sustained immunity with experimental EBOV vaccines in NHPs remains unknown; however, in a recent study, a decrease in GP-specific IgG levels due to old age or a suboptimal reaction to the VSVΔG/EBOVGP vaccine in rodents also seem to be indicative of non-survival<sup>28</sup>.

ZMapp consists of a cocktail of highly purified monoclonal antibodies; which constitutes a less controversial alternative than whole-blood transfusions from convalescent survivors, as was performed during the 1995 EBOV outbreak in Kikwit<sup>29</sup>. The safety of monoclonal antibody therapy is well documented, with generally low rates of adverse reactions, the capacity to confer rapid and specific immunity in all populations, including the young, the elderly and the immunocompromised, and if necessary, the ability to provide higher-than-natural levels of immunity compared to vaccinations<sup>30</sup>. The evidence presented here suggests that ZMapp offers the best option of the experimental therapeutics currently in development for treating EBOV-infected patients. We hope that initial safety testing in humans will be undertaken soon, preferably within the next few months, to enable the compassionate use of ZMapp as soon as possible.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 5 August; accepted 21 August 2014.**

**Published online 29 August 2014.**

- Bausch, D. G., Sprecher, A. G., Jeffs, B. & Boumandouki, P. Treatment of Marburg and Ebola hemorrhagic fevers: a strategy for testing new drugs and vaccines under outbreak conditions. *Antiviral Res.* **78**, 150–161 (2008).
- Baize, S. *et al.* Emergence of Zaire Ebola virus disease in Guinea — preliminary report. *N. Engl. J. Med.* <http://dx.doi.org/10.1056/NEJMoa1404505> (2014).
- WHO. Ebola virus disease (EVD) <http://www.who.int/csr/don/archive/disease/ebola/en/> (accessed, 15 August 2014).
- CDC. Chronology of Ebola Hemorrhagic Fever Outbreaks, <http://www.cdc.gov/vhf/ebola/resources/outbreak-table.html> (accessed, 15 August 2014).
- Reliefweb. W. African Ebola epidemic 'likely to last months': UN <http://reliefweb.int/report/guinea/w-african-ebola-epidemic-likely-last-months-un> (7 March 2014).
- Clark, D. V., Jahrling, P. B. & Lawler, J. V. Clinical management of filovirus-infected patients. *Viruses* **4**, 1668–1686 (2012).
- Guimard, Y. *et al.* Organization of patient care during the Ebola hemorrhagic fever epidemic in Kikwit, Democratic Republic of the Congo, 1995. *J. Infect. Dis.* **179** (Suppl. 1), S268–S273 (1999).
- Hensley, L. E. *et al.* Recombinant human activated protein C for the postexposure treatment of Ebola hemorrhagic fever. *J. Infect. Dis.* **196** (Suppl 2), S390–S399 (2007).

- Geisbert, T. W. *et al.* Treatment of Ebola virus infection with a recombinant inhibitor of factor VIIa/tissue factor: a study in rhesus monkeys. *Lancet* **362**, 1953–1958 (2003).
- Geisbert, T. W. *et al.* Postexposure protection of non-human primates against a lethal Ebola virus challenge with RNA interference: a proof-of-concept study. *Lancet* **375**, 1896–1905 (2010).
- Warren, T. K. *et al.* Advanced antisense therapies for postexposure protection against lethal filovirus infections. *Nature Med.* **16**, 991–994 (2010).
- Feldmann, H. *et al.* Effective post-exposure treatment of Ebola infection. *PLoS Pathog.* **3**, e2 (2007).
- Olinger, G. G., Jr *et al.* Delayed treatment of Ebola virus infection with plant-derived monoclonal antibodies provides protection in rhesus macaques. *Proc. Natl Acad. Sci. USA* **109**, 18030–18035 (2012).
- Qiu, X. *et al.* Successful treatment of ebola virus-infected cynomolgus macaques with monoclonal antibodies. *Sci. Transl. Med.* **4**, 138ra181 (2012).
- Pettitt, J. *et al.* Therapeutic intervention of Ebola virus infection in rhesus macaques with the MB-003 monoclonal antibody cocktail. *Sci. Transl. Med.* **5**, 199ra113 (2013).
- Qiu, X. *et al.* mAbs and Ad-vectored IFN-α therapy rescue Ebola-infected nonhuman primates when administered after the detection of viremia and symptoms. *Sci. Transl. Med.* **5**, 207ra143 (2013).
- Giritch, A. *et al.* Rapid high-yield expression of full-size IgG antibodies in plants coinfecting with noncompeting viral vectors. *Proc. Natl Acad. Sci. USA* **103**, 14701–14706 (2006).
- Jahrling, P. B. *et al.* Evaluation of immune globulin and recombinant interferon-α2b for treatment of experimental Ebola virus infections. *J. Infect. Dis.* **179** (Suppl 1), S224–S234 (1999).
- Qiu, X. *et al.* Ebola GP-specific monoclonal antibodies protect mice and guinea pigs from lethal Ebola virus infection. *PLoS Negl. Trop. Dis.* **6**, e1575 (2012).
- Wilson, J. A. *et al.* Epitopes involved in antibody-mediated protection from Ebola virus. *Science* **287**, 1664–1666 (2000).
- Qiu, X. *et al.* Characterization of Zaire ebolavirus glycoprotein-specific monoclonal antibodies. *Clin. Immunol.* **141**, 218–227 (2011).
- Dye, J. M. *et al.* Postexposure antibody prophylaxis protects nonhuman primates from filovirus disease. *Proc. Natl Acad. Sci. USA* **109**, 5034–5039 (2012).
- Wong, G. *et al.* Immune parameters correlate with protection against ebola virus infection in rodents and nonhuman primates. *Sci. Transl. Med.* **4**, 158ra146 (2012).
- Marzi, A. *et al.* Antibodies are necessary for rVSV/ZEBV-GP-mediated protection against lethal Ebola virus challenge in nonhuman primates. *Proc. Natl Acad. Sci. USA* **110**, 1893–1898 (2013).
- Promedmail.org. Ebola virus disease - West Africa (117): WHO, Nigeria, Liberia, drug, more. <http://www.promedmail.org/direct.php?id=2666073> (6 August 2014).
- NC3RS. Practical blood sample volumes for laboratory animals, domestic species and non-human primates. <http://www.nc3rs.org.uk/bloodsamplingmicrosite/page.asp?id=426> (accessed, 3 August 2014).
- Qiu, X. *et al.* Sustained protection against Ebola virus infection following treatment of infected nonhuman primates with ZMAB. *Sci. Rep.* **3**, 3365 (2013).
- Wong, G. *et al.* Immunization with vesicular stomatitis virus vaccine expressing the Ebola glycoprotein provides sustained long-term protection in rodents. *Vaccine* <http://dx.doi.org/10.1016/j.vaccine.2014.08.028> (in the press).
- Mupapa, K. *et al.* Treatment of Ebola hemorrhagic fever with blood transfusions from convalescent patients. International Scientific and Technical Committee. *J. Infect. Dis.* **179** (Suppl. 1), S18–S23 (1999).
- UPMHealthsecurity.org. Next-Generation Monoclonal Antibodies: Challenges and Opportunities [http://www.upmhealthsecurity.org/our-work/pubs\\_archive/pubs-pdfs/2013/2013-02-04-next-gen-monoclonal-antibodies.pdf](http://www.upmhealthsecurity.org/our-work/pubs_archive/pubs-pdfs/2013/2013-02-04-next-gen-monoclonal-antibodies.pdf) (UPMC Center for Biosecurity, 2013).

**Acknowledgements** The authors thank K. Tierney, A. Grolla, S. Jones, J. Dong and D. Kobasa for their technical assistance, V. Klimyuk and Y. Gleba for access to the magnICON expression system, and H. Steinkellner for access to transgenic *N. benthamiana*. This work was supported by the Defense Threat Reduction Agency (DTRA contract HDTRA1-13-C-0018), the National Institutes of Health (U19AI109762), the Public Health Agency of Canada (PHAC), and a Canadian Safety and Security Program (CSSP) grant to G.P.K. and X.Q. G.W. is the recipient of a Doctoral Research Award from the Canadian Institute for Health Research (CIHR).

**Author Contributions** X.Q., G.P.K. and L.Z. designed the experiments. X.Q., G.W., J.A., A.B., L.F., J.B.A., H.F., H.W., J.A., J. P., G.G.O. and G.P.K. performed the experiments. X.Q., G.W., J.A., K.W., B.X., J.E.S., L.Z. and G.P.K. wrote the manuscript. E.H., A.J., J.M., K.S., O.B., N.B., C.G., D.K., M.H.P., J.V., K.W. and L.Z. contributed reagents for this study.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.P.K. ([gary.kobinger@phac-aspc.gc.ca](mailto:gary.kobinger@phac-aspc.gc.ca)) or L.Z. ([larry.zeitlin@mappbio.com](mailto:larry.zeitlin@mappbio.com)).



## METHODS

**Ethics statement.** The guinea pig experiment, in addition to the second and third NHP study (ZMapp1, ZMapp2 and ZMapp) were performed at the National Microbiology Laboratory (NML) as described on Animal Use Document (AUD) #H-13-003, and has been approved by the Animal Care Committee (ACC) at the Canadian Science Center for Human and Animal Health (CSCAH), in accordance with the guidelines outlined by the Canadian Council on Animal Care (CCAC). The first study with MB-003 in NHPs was performed at United States Army Medical Research Institute of Infectious Diseases (USAMRIID) under an Institutional Animal Care and Use Committee (IACUC) approved protocol in compliance with the Animal Welfare Act, Public Health Service Policy, and other federal statutes and regulations relating to animals and experiments involving animals. The facility where this research was conducted is accredited by The Association for Assessment and Accreditation of Laboratory Animal Care International and adheres to principles stated in the 8th edition of the *Guide for the Care and Use of Laboratory Animals*, National Research Council (2011; <http://grants.nih.gov/grants/olaw/Guide-for-the-care-and-use-of-laboratory-animals.pdf>).

**Monoclonal antibody production.** The large-scale production of mAb cocktails cZMAb, MB-003, ZMapp1, ZMapp2 and ZMapp in addition to control mAb 4E10 (anti-HIV) from *N. benthamiana* under GMP conditions was done by Kentucky BioProcessing (Owensboro, KY) as described previously<sup>13,15,31</sup>. The large-scale production of m4G7 was performed by the National Research Council (Montreal site) using a previously described protocol<sup>16</sup>.

**Viruses.** The challenge virus used in NHPs was Ebola virus *H.sapiens-tc/COD/1995/Kikwit-9510621* (EBOV-K) (order *Mononegavirales*, family *Filoviridae*, species *Zaire ebolavirus*; GenBank accession no. AY354458)<sup>18</sup>. Passage three from the original stock was used for the studies at the NML and passage four was used for the study performed at USAMRIID (the NHP study with the individual MB-003 mAbs). Sequencing of 112 clones from the passage three stock virus revealed that the population ratio of 7U:8U in the EBOV GP editing site was 80:20; sequencing for the passage four stock virus was not performed, and therefore the ratio of 7U:8U in the editing site was unknown. The virus used in guinea pig studies was guinea-pig-adapted EBOV, Ebola virus VECTOR/C.porcus-lab/COD/1976/Mayinga-GPA (EBOV-M-GPA) (order *Mononegavirales*, family *Filoviridae*, species *Zaire ebolavirus*; GenBank accession number AF272001.1)<sup>32</sup>. The Guinean variant used in IgG ELISA and neutralizing antibody assays was Ebola virus *H.sapiens-tc/GIN/2014/Guekedou-C05* (EBOV-G) (order *Mononegavirales*, family *Filoviridae*, species *Zaire ebolavirus*; GenBank accession no. KJ660348.1)<sup>2</sup>.

**Animals.** Outbred 6–8-week-old female Hartley strain guinea pigs (Charles River) were used for these studies. Animals were infected intraperitoneally with  $1,000 \times \text{LD}_{50}$  of EBOV-M-GPA. The animals were then treated with one dose of ZMAb, MB-003, ZMapp1, ZMapp2, c13C6, h13F6 or c6D8 totalling 5 mg per guinea pig, and monitored every day for 28 days for survival, weight and clinical symptoms. This study was not blinded, and no animals were excluded from the analysis.

For the MB-003 study performed at USAMRIID, thirteen rhesus macaques (*Macaca mulatta*) were obtained from the USAMRIID primate holding facility, ranging from 5.1 to 10 kg. This study was not blinded, and no animals were excluded from the analysis. Animals were given standard monkey chow, primate treats, fruits, and vegetables for the duration of the study. All animals were challenged intramuscularly with a target dose of  $1,000 \text{ p.f.u.}$  Treatment with either monoclonal antibody, MB-003 cocktail, or PBS was administered on 1, 4, and 7 dpi via saphenous intravenous infusion. Animals were monitored at least once daily for changes in health, diet, behaviour, and appearance. Animals were sampled for chemical analysis, complete bloods counts and viraemia on 0, 3, 5, 7, 10, 14, 21, and 28 dpi.

For the ZMapp1 and ZMapp2 study, fourteen male and female rhesus macaques (*Macaca mulatta*), ranging from 4.1 to 9.6 kg (4–8 years old) were purchased from Primen (USA). This study was not blinded, and no animals were excluded from the analysis. Animals were assigned groups based on gender and weight. Animals were fed standard monkey chow, fruits, vegetables, and treats. Husbandry enrichment consisted of visual stimulation and commercial toys. All animals were challenged intramuscularly with a high dose of EBOV (backtitre:  $4,000 \times \text{TCID}_{50}$  or  $2,512 \text{ p.f.u.}$ ) at 0 dpi. Administration of the first treatment dose was initiated at 3 dpi, with identical doses at 6 and 9 dpi. Animals were scored daily for signs of disease, in addition to changes in food and water consumption. On designated treatment days in addition to 14, 21, and 27 dpi, the rectal temperature and clinical score were measured, and the following were sampled: blood for serum biochemistry and cell counts and viraemia. This study was not blinded, and no animals were excluded from the analysis.

For the ZMapp study, twenty-one male rhesus macaques, ranging from 2.5 to 3.5 kg (2 years old) were purchased from Primen (USA). This study was not blinded, and no animals were excluded from the analysis. Animals were assigned groups based on gender and weight. Animals were fed standard monkey chow, fruits, vegetables, and treats. Husbandry enrichment consisted of visual stimulation and commercial toys. All animals were challenged intramuscularly with EBOV (backtitre:  $1,000 \times \text{TCID}_{50}$  or  $628 \text{ p.f.u.}$ ) at 0 dpi. Administration of the first treatment dose was initiated at 3, 4 or 5 dpi, with two additional identical doses spaced 3 days apart. Animals were scored daily for signs of disease, in addition to changes in food and water consumption. On designated treatment days in addition to 14, 21, and 28 dpi, the rectal temperature and clinical score were measured, and the following were sampled: blood for serum biochemistry and cell counts and viraemia.

**Blood counts and blood biochemistry.** Complete blood counts were performed with the VetScan HM5 (Abaxis Veterinary Diagnostics). The following parameters were shown in the figures: levels of white blood cells, lymphocytes, percentage of lymphocytes, levels of platelets, neutrophils and percentage of neutrophils. Blood biochemistry was performed with the VetScan VS2 (Abaxis Veterinary Diagnostics). The following parameters were shown in the figures: levels of alkaline phosphatase, alanine aminotransferase, blood urea nitrogen, creatinine, total bilirubin and glucose. **Enzyme-linked immunosorbent assays (ELISAs).** IgG ELISA with c13C6, c2G4 or c1H3 was performed as described previously<sup>16</sup> using gamma-irradiated EBOV-G and EBOV-K virions purified on a 20% sucrose cushion as the capture antigen in the ELISA. Each mAb was assayed in triplicate.

**Neutralizing antibody assays.** Twofold dilutions of c13C6, c2G4 or c1H3 ranging from 0.0156 to  $2 \mu\text{g ml}^{-1}$  were first incubated with  $100 \text{ p.f.u.}$  of EBOV-G at room temperature for 1 h with or without complement, transferred to Vero E6 cells and incubated at  $37^\circ\text{C}$  for 1 h, and then replaced with DMEM supplemented with 2% fetal bovine serum and scored for the presence of cytopathic effect at 14 dpi. The lowest concentrations of mAbs demonstrating the absence of cytopathic effect were averaged and reported.

**EBOV titration by  $\text{TCID}_{50}$  and RT-qPCR.** Titration of live EBOV was determined by adding tenfold serial dilutions of whole blood to VeroE6 cells, with three replicates per dilution. The plates were scored for cytopathic effect at 14 dpi, and titres were calculated with the Reed and Muench method<sup>33</sup>. Results were shown as median tissue culture infectious dose ( $\text{TCID}_{50}$ ).

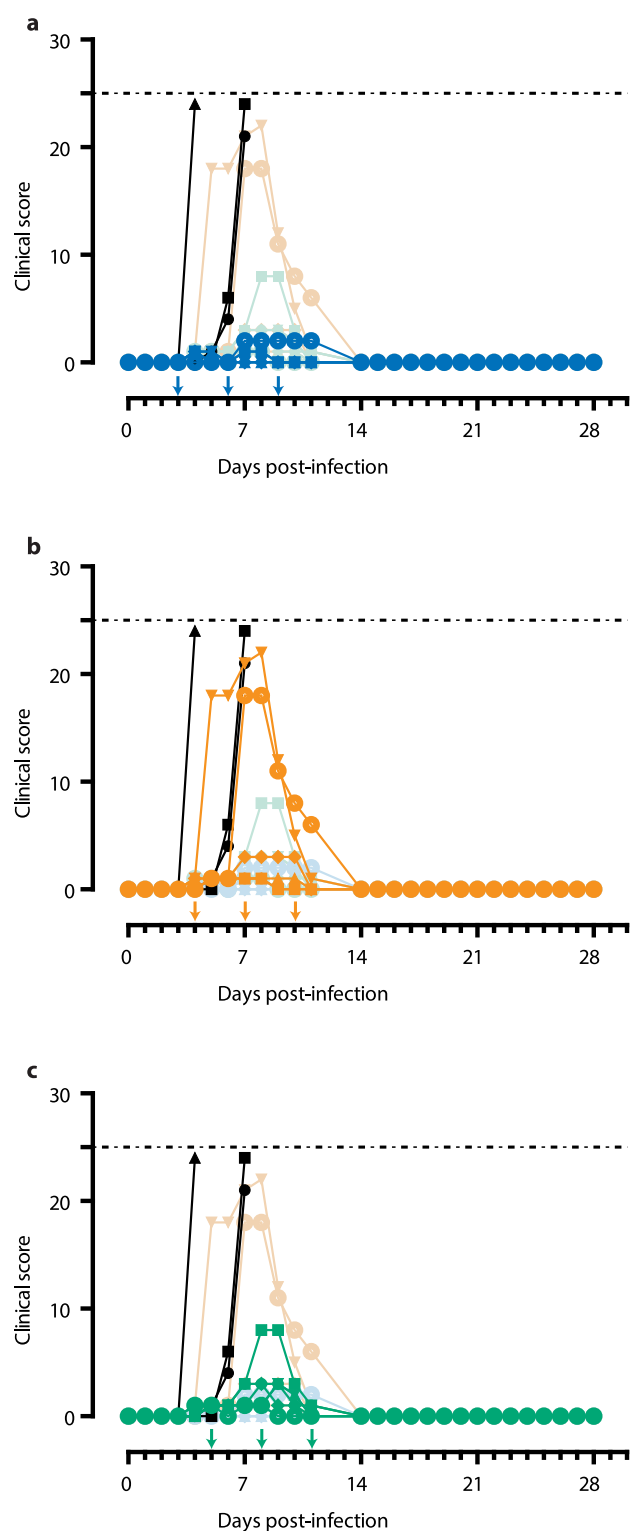
For titres measured by RT-qPCR, total RNA was extracted from whole blood with the QIAmp Viral RNA Mini Kit (Qiagen). EBOV was detected with the LightCycler 480 RNA Master Hydrolysis Probes (Roche) kit, with the RNA polymerase (nucleotides 16472 to 16538, AF086833) as the target gene. The reaction conditions were as follows:  $63^\circ\text{C}$  for 3 min,  $95^\circ\text{C}$  for 30 s, and cycling of  $95^\circ\text{C}$  for 15 s,  $60^\circ\text{C}$  for 30 s for 45 cycles on the ABI StepOnePlus. The lower detection limit for this assay is 86 genome equivalents  $\text{ml}^{-1}$ . The sequences of primers used were as follows: EBOVLF2 (CAGCCAGCAATTTCTTCCAT), EBOVLR2 (TTTCGGTTGCTGTTTCTGTG), and EBOVLP2FAM (FAM-ATCATTGGCGTACTGGAGGAGCAG-BHQ1).

**Sequence alignment.** Protein sequences for EBOV-K and EBOV-G surface glycoproteins were obtained from GenBank, accession numbers AGB56794.1 and AHX24667.1 respectively. The sequences were aligned using DNASTAR Lasergene 10 MEGAlign using the Clustal W algorithm.

**Statistical analysis.** For the guinea pig and nonhuman primate studies, each treatment group consisted of six animals. Assuming a significance threshold of 0.05, a sample size of six per group will give  $>80\%$  power to detect a difference in survival proportions between the treatment (83% survival or higher) and the control group using a one-tailed Fisher's exact test.

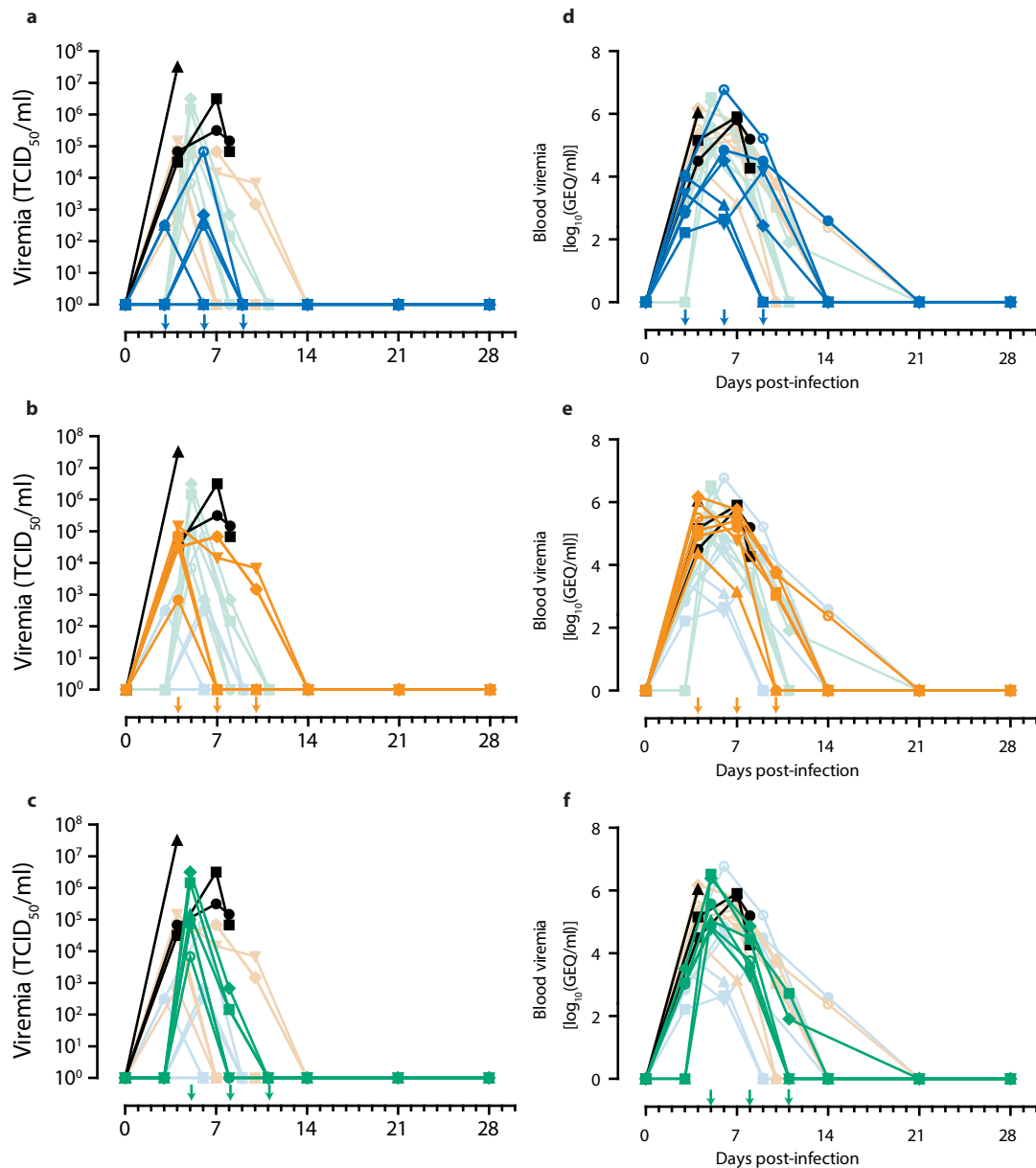
Survival was compared using the log-rank test in GraphPad PRISM 5, differences in survival were considered significant when the *P* value was less than 0.05. Antibody binding to EBOV-G and EBOV-K was compared by fitting the data to a four-parameter logistic regression using GraphPad PRISM 5. The  $\text{EC}_{50}$  were considered different if the 95% confidence intervals excluded each other. For all statistical analyses, the data conformed to the assumptions of the test used.

- Zeitlin, L. *et al.* Enhanced potency of a fucose-free monoclonal antibody being developed as an Ebola virus immunoprotectant. *Proc. Natl Acad. Sci. USA* **108**, 20690–20694 (2011).
- Connolly, B. M. *et al.* Pathogenesis of experimental Ebola virus infection in guinea pigs. *J. Infect. Dis.* **179** (Suppl. 1), S203–S217 (1999).
- Reed, L. J. & Muench, H. A simple method of estimating fifty per cent endpoints. *Am. J. Hyg.* **27**, 493–497 (1938).



**Extended Data Figure 1 | Clinical scores for each ZMapp-treated group.** Arrows indicate treatment days. Dashed line represents humane endpoint threshold. Faded symbols/lines are the other two treatment groups, for comparison. Control group (Group G) is shown in black on all three panels. **a**, Clinical score of Group D (blue); **b**, clinical score of Group E (orange); **c**, clinical score of Group F (green).





**Extended Data Figure 2 | Viraemia for each ZMapp-treated group.** Arrows indicate treatment days. Faded symbols/lines are the other two treatment groups, for comparison. Control group (Group G) is shown in black on all three

panels. **a**, TCID<sub>50</sub> of Group D (blue); **b**, TCID<sub>50</sub> of Group E (orange); **c**, TCID<sub>50</sub> of Group F (green). **d**, Viraemia by RT-qPCR of Group D (blue); **e**, viraemia by RT-qPCR of Group E (orange); **f**, viraemia by RT-qPCR of Group F (green).

**Extended Data Table 1 | Blood viraemia measured by RT-qPCR for the ZMapp1- and ZMapp2-treated NHPs**

Day	A1	A2	A3	A4	A5	A6	B1	B2	B3	B4	B5	B6	C1	C2
0	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD
3	UD	3.98E+02	UD	UD	9.99E+02	1.27E+03	8.05E+03	1.65E+04	9.36E+03	9.77E+03	9.27E+02	9.48E+02	UD	4.34E+02
6	3.10E+03	4.49E+02	UD	8.34E+02	5.81E+03	2.09E+03	UD	1.22E+04	1.04E+05	4.26E+03	3.14E+02	4.49E+03	5.57E+06	2.05E+07
7													5.50E+05	
9		UD	UD	UD	UD	5.24E+02	UD	1.74E+05	5.03E+05	1.87E+03	5.16E+02	UD		
14	3.62E+03	UD	UD	UD	UD	UD	UD	UD		UD	UD	UD		
21	UD	UD	UD	UD	UD	UD	UD	UD		UD	UD	UD		
27	UD	UD	UD	UD	UD	UD	UD	UD		UD	UD	UD		

UD, undetectable.



Extended Data Table 2 | Oral swab viraemia measured by RT–qPCR for the ZMapp1- and ZMapp2-treated NHPs

Days	A1	A2	A3	A4	A5	A6	B1	B2	B3	B4	B5	B6	C1	C2
0	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD
3	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD
6	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD
7													5.05E+03	
9		UD	UD	UD	UD	UD	UD	UD	4.81E+04	UD	UD	UD		
14	UD	UD	UD	UD	UD	UD	UD	UD		UD	UD	UD		
21	UD	UD	UD	UD	UD	UD	UD	UD		UD	UD	UD		
27	UD	UD	UD	UD	UD	UD	UD	UD		UD	UD	UD		

UD, undetectable.

**Extended Data Table 3 | Nasal swab viraemia measured by RT-qPCR for the ZMapp1- and ZMapp2-treated NHPs**

Days	A1	A2	A3	A4	A5	A6	B1	B2	B3	B4	B5	B6	C1	C2
0	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD
3	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD
6	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	3.75E+02
7													1.98E+04	2.16E+03
9		UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD		
14	UD	UD	UD	UD	UD	UD	UD	UD		UD	UD	UD		
21	UD	UD	UD	UD	UD	UD	UD	UD		UD	UD	UD		
27	UD	UD	UD	UD	UD	UD	UD	UD		UD	UD	UD		

UD, undetectable.



Extended Data Table 4 | Rectal swab viraemia measured by RT–qPCR for the ZMapp1- and ZMapp2-treated NHPs

Days	A1	A2	A3	A4	A5	A6	B1	B2	B3	B4	B5	B6	C1	C2
0	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD
3	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD
6	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	UD	4.16E+02	8.17E+03
7													4.38E+04	
9		UD	UD	UD	UD	UD	UD	UD	3.90E+02	UD	UD	UD		
14	UD	UD	UD	UD	UD	UD	UD	UD		UD	UD	UD		
21	UD	UD	UD	UD	UD	UD	UD	UD		UD	UD	UD		
27	UD	UD	UD	UD	UD	UD	UD	UD		UD	UD	UD		

UD, undetectable.

**Extended Data Table 5 | Blood viraemia measured by RT-qPCR for the ZMapp-treated NHPs**

Days	A1	A2	A3	A4	A5	A6	B1	B2	B3	B4	B5	B6	C1	C2	C3	C4	C5	C6	D1	D2	D3
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	676.08	165.96	10233	2884	812.83	10965							1047.1			1122	3235.9	1148.2			
4							85114	128825	23442	1E+06	1E+06	323594							31623	144544	1E+06
5													380189	3E+06	109648	58884	2E+06	69183			
6	70795	446.68	1230.3	316.23	32359	6E+06															
7							154882	257040	1380.4	63096	588844	363078							645654	812831	
8													3715.4	28184	29512	1862.1	72444	5888.4	158489	18621	
9	31623	1	1	15136	275.42	165959															
10							1	1071.5	1	1318.3	6166	5248.1									
11													1	524.81	1	1	81.283	1			
14	398.11	1	1	1	1	1	1	1	1	1	1	239.88	1	1	1	1	1	1			
21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1			
28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1			



# Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity

Andriy Marusyk<sup>1,2,3</sup>, Doris P. Tabassum<sup>1,4</sup>, Philipp M. Altrock<sup>5,6,7</sup>, Vanessa Almendro<sup>1,2,3</sup>, Franziska Michor<sup>5,6</sup> & Kornelia Polyak<sup>1,2,3,4,8</sup>

**Cancers arise through a process of somatic evolution that can result in substantial sub-clonal heterogeneity within tumours. The mechanisms responsible for the coexistence of distinct sub-clones and the biological consequences of this coexistence remain poorly understood. Here we used a mouse xenograft model to investigate the impact of sub-clonal heterogeneity on tumour phenotypes and the competitive expansion of individual clones. We found that tumour growth can be driven by a minor cell subpopulation, which enhances the proliferation of all cells within a tumour by overcoming environmental constraints and yet can be outcompeted by faster proliferating competitors, resulting in tumour collapse. We developed a mathematical modelling framework to identify the rules underlying the generation of intra-tumour clonal heterogeneity. We found that non-cell-autonomous driving of tumour growth, together with clonal interference, stabilizes sub-clonal heterogeneity, thereby enabling inter-clonal interactions that can lead to new phenotypic traits.**

Cancers result from genetic and epigenetic changes that fuel Darwinian somatic evolution<sup>1,2</sup>. Until recently, the evolution was assumed to proceed as a linear succession of clonal expansions triggered by acquisition of strong driver mutations that progressively increase cell fitness and lead to selective sweeps<sup>3</sup>. However, recent data from tumour genome sequencing studies and single-cell based analyses has revealed substantial genetic heterogeneity within tumours, including sub-clonal differences in driver mutations<sup>4–8</sup>. This contradicts the linear succession model and challenges the assumption of tumour evolution being driven by mutations providing strong clone-specific selective advantages. Furthermore, clonal heterogeneity raises the possibility of biologically and clinically important interactions between distinct clones<sup>9,10</sup>.

Many oncogenic mutations confer a cell-autonomous fitness advantage by either providing independence from growth factors or abolishing an apoptotic response. These mutations are thus expected to drive clonal expansions<sup>11</sup>. At the same time, tumour progression is frequently limited by microenvironmental constraints<sup>12–14</sup> that cannot be overcome by a cell-autonomous increase in proliferation rates. Instead, progression depends on alterations of the microenvironment, mediated by factors acting non-cell-autonomously, such as metalloproteinases and cytokines. It is unclear whether these secreted factors preferentially benefit the ‘producer’ clone(s) enabling their clonal dominance.

## A model of clonal heterogeneity

Understanding clonal heterogeneity has been hindered by the lack of suitable experimental models. Although patient tumour-derived xenograft studies using clonal tracing can be insightful, their utility is limited by the challenges in deciphering mechanisms that underlie biological differences between sub-clones. We aimed to bypass these challenges by experimentally defining sub-populations via overexpression of factors previously implicated in tumour progression. We decided to exploit a scenario of a tumour that is ‘stuck’ in a microenvironmentally constrained progression bottleneck, which is relevant for clinically asymptomatic cancers, dormant micro-metastatic lesions and perhaps early clinically

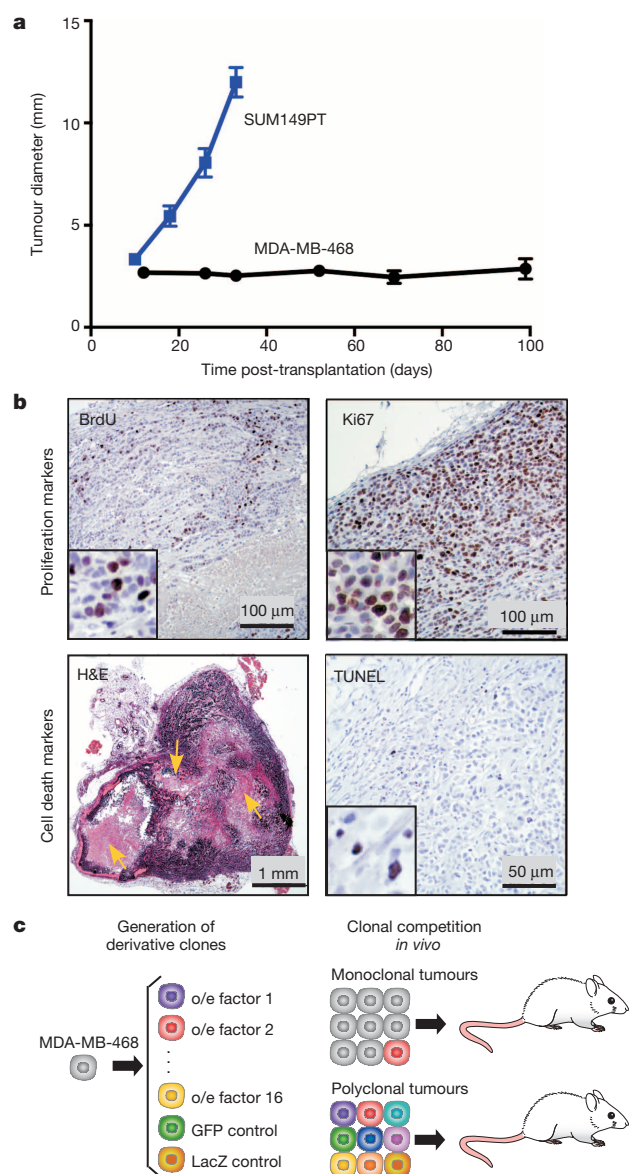
undetectable stages of tumour development. This scenario offers two key advantages. First, in contrast to a rapidly growing tumour, the constrained population size of non-growing tumours composed of rapidly cycling cells is expected to intensify competition for limited microenvironmental resources. This enhances the detection of differences in competitive fitness. Second, the indolent morphology and lack of net tumour growth should facilitate the detection of increase in tumour growth and metastasis.

In search of tumours satisfying these criteria, we analysed a panel of breast cancer-derived cell lines for tumours formed by orthotopic transplantation into the mammary fat pads of immunodeficient *Foxn1<sup>nu</sup>* (nu) mice. Whereas most of the tested cell lines either failed to produce tumours or formed tumours that grew too rapidly (for example, SUM149PT cells), the MDA-MB-468 cell line formed indolent tumours which, upon reaching 2–5 mm in diameter, showed very slow growth rates (Fig. 1a and data not shown). Despite slow net growth, the tumour cells were actively proliferating: 80–90% of them were in the cell cycle based on Ki-67 staining, and 20–30% were in S phase based on 5-bromodeoxyuridine (BrdU) incorporation (Fig. 1b). The slow net tumour growth indicated that cell proliferation was counterbalanced by cell death. Indeed, 1–3% of the cells were apoptotic. Tumours contained large necrotic areas indicating substantial necrotic cell death (Fig. 1b).

We used MDA-MB-468 cells to generate a panel of sub-lines (henceforth called ‘sub-clones’) defined by the lentiviral overexpression of a single secreted factor. Each factor had been previously implicated in tumour progression, along with reported high expression in breast carcinoma cells of patients (Fig. 1c and Extended Data Table 1). Given the recently reported variability in clonal proliferation dynamics<sup>15</sup> and to minimize the confounding influences of genetic/epigenetic heterogeneity within the cell lines, we used pools of transduced cells rather than single cell-derived clones. This panel enabled us to compare phenotypic properties of tumours and clonal expansions under two circumstances: (1) each sub-clone competing against parental cells (monoclonal tumours), and (2) sub-clones competing against all other sub-clones (polyclonal tumours)

<sup>1</sup>Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02215, USA. <sup>2</sup>Department of Medicine, Brigham and Women’s Hospital, Boston, Massachusetts 02115, USA.

<sup>3</sup>Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>4</sup>BBS Program, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>5</sup>Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, Massachusetts 02215, USA. <sup>6</sup>Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts 02115, USA. <sup>7</sup>Program for Evolutionary Dynamics, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>8</sup>Harvard Stem Cell Institute and the Broad Institute, Cambridge, Massachusetts 02138, USA.



**Figure 1 | Experimental system.** **a**, Growth of tumours upon mammary fat pad transplantation of indicated cell lines,  $n = 10$  per group, combined data from 2 independent experiments, error bars indicate s.e.m.

**b**, Representative images of indicated staining. Arrows indicate necrotic areas. H&E, haematoxylin and eosin. **c**, Experimental scheme.

(Fig. 1c). We had 18 sub-clones in total. In order to maintain equal initial clonal proportions in all tumours, we employed the cell number ratio of 1:18 between a sub-clone and parental competitors.

### Non-cell-autonomous tumour driving

We first investigated whether individual sub-clones, initially present as a minor sub-population competing against parental cells, could affect tumour properties. We focused on tumour growth and metastasis, features that are most relevant clinically and amenable to quantification. Although we observed variability between the groups in morphology, proliferation and vascularization (Extended Data Fig. 1), only the chemokine (C-C motif) ligand 5 (CCL5) and interleukin 11 (IL11) overexpressing sub-clones were able to enhance tumour growth (Fig. 2a, b). None of the tumours were metastatic, as evaluated by *in vivo* bioluminescence imaging and examination of draining lymph nodes, peritoneal walls and bone marrow (data not shown).

We then analysed the population frequencies of individual sub-clones within the tumours using a genomic quantitative polymerase chain reaction

(qPCR) approach, using clone-specific and reference amplicons (Extended Data Fig. 2). Surprisingly, we observed no strict correlation between the increase in sub-clonal frequencies and the growth rate of tumours (Fig. 2a–c). The LOXL3-overexpressing sub-clone underwent the greatest (~tenfold) expansion in population frequency, yet failed to promote overall tumour growth. On the other hand, both CCL5 and IL11, each capable of driving outgrowth of tumours, exhibited approximately eightfold and fourfold expansion, respectively. To address the link between clone-specific expansion and tumour growth more directly, we calculated rates of expansion in cell numbers over the initially transplanted cells using a volume-based cellularity inference of  $4.1 \times 10^5$  cells per  $\text{mm}^3$  (Fig. 2d, Extended Data Fig. 3a). Only IL11 was capable of non-cell-autonomous tumour growth driving. We saw enhanced expansion of both IL11-expressing and parental cells. Increased growth of CCL5-driven tumours was only attributable to cell-autonomous expansion of CCL5-expressing cells. This finding was consistent with the observed delay in tumour outgrowth driven by CCL5 compared to IL11-driven tumours (Fig. 2a, inset).

We did not observe a positive correlation between tumour weights and final percentages of IL11 expressing cells (Extended Data Fig. 4a). An increase in the initial frequency of the IL11 sub-clone also did not further enhance tumour growth (Extended Data Fig. 4b). Parental cells expressed undetectable basal levels of IL11 (Extended Data Fig. 4c, d) and the non-cell-autonomous driving of tumour growth was observed with four independent derivations of the IL11 overexpressing sub-clones using two distinct lentiviral backbones that provide different levels of expression (Extended Data Fig. 4c–e). This observation strongly suggests that the phenomenon was IL11-specific and did not require additional stochastic events.

We then initiated tumours in which all the sub-clones, present at the initial 1:18 ratio, were set to compete against one another. These tumours grew faster than monoclonal tumours, suggesting additive growth-promoting interactions among the sub-clones (Fig. 2a). However, omitting the IL11 sub-clone (2:18 ratio of control LacZ sub-clone was used to maintain 1:18 ratio of the remaining sub-clones) blocked the increased growth of polyclonal tumours, reducing clonal expansions (Fig. 2e and Extended Data Fig. 5a). Therefore, non-cell-autonomous stimulation by IL11 was both necessary and sufficient to drive tumour growth.

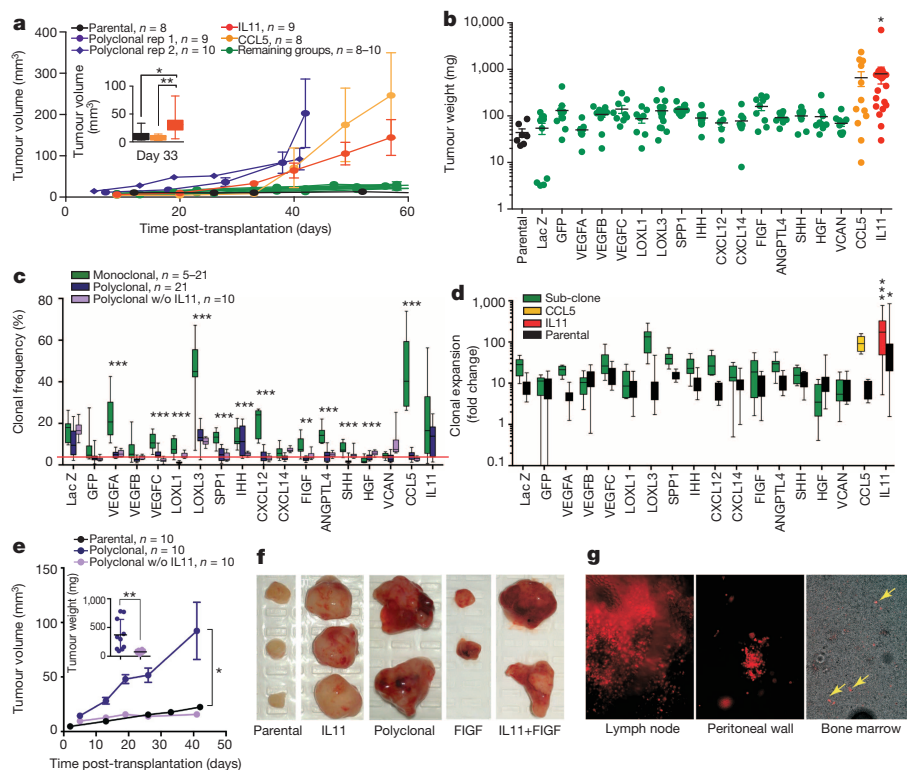
### Sub-clonal cooperation in metastasis

In addition to accelerated growth rates, polyclonal tumours displayed regions of extensive haemorrhage and multiple cysts (Fig. 2f), indicative of increased blood and lymphatic vessel leakage. Consistently, a large fraction of polyclonal tumours were metastatic: 7/12 analysed animals displayed lymph node metastases, 6/12 displayed metastatic nodes on the peritoneal wall and 4/7 contained tumour cells in the bone marrow (Fig. 2g). Animals bearing polyclonal tumours accumulated peritoneal fluid and demonstrated signs of systemic toxicity, requiring euthanasia at earlier time points compared to other groups.

FIGF was the only other sub-clone displaying elevated vascular leakage in monoclonal tumours, albeit with incomplete penetrance. Hence we asked whether the combination of IL11 and FIGF could recapitulate the metastatic phenotypes of polyclonal tumours. Indeed, FIGF/IL11 tumours displayed an increase in tumour volume and extensive haemorrhage (Fig. 2f, Extended Data Fig. 5b), with 4/4 animals presenting both lymph node and peritoneal wall metastases. Therefore, our data suggest that biological interactions between distinct sub-populations can lead to the emergence of new tumour phenotypes.

### Mechanisms of IL11-driven tumour growth

Elevated tumour growth implies an increase in net cell proliferation rates, either by stimulating proliferation or by inhibiting cell death. IL11-driven tumours displayed a subtle, but significant, increase in proliferation rates compared to parental tumours (Fig. 3a). Apoptosis rates were similar (Extended Data Fig. 1b). This increase in cellular proliferation could result either from a direct autocrine/paracrine stimulation of cell



**Figure 2 | Polyclonality affects tumour phenotypes.** **a**, Tumour growth kinetics. **b**, Tumour weights. **c**, Sub-clones frequencies within tumours. Red line indicates initial frequency. **d**, Expansion (fold-change over initial cell number) of sub-clones and parental cells from monoclonal tumours shown in **c**. **e**, Tumour growth kinetics and weights (inset). **f**, Representative images of tumours. **g**, live fluorescent microscopy images of tumour cells (mCherry<sup>+</sup>) in tissues. \* $P < 0.05$ , \*\* $P < 0.01$  and \*\*\* $P < 0.001$ , respectively, of Student's *t*-test (**a**, **c**, **e**) or ANOVA multiple group comparison against parental (**b**) or LacZ (**d**) with Dunnett's correction. Error bars indicate s.e.m. Data shown are representative of at least 2 independent experiments.

growth or from indirect effects mediated by the microenvironment. IL11 signals through a unique and specific receptor, IL11R $\alpha$ , that forms a signalling complex with the GP130 co-receptor shared with other IL6 cytokine family members<sup>17</sup>. IL11 promotes growth of gastric carcinoma via direct stimulation of epithelial cells<sup>18,19</sup>. Similar stimulation of tumour growth via non-cell-autonomous signalling between tumour cells, involving two related cytokines, IL6 and LIF, was reported in glioblastomas<sup>20</sup>. We therefore asked whether modulation of IL11R $\alpha$  expression in carcinoma cells affects the ability of IL11 to induce tumour growth. Neither overexpression nor short hairpin (shRNA)-mediated downregulation of IL11R $\alpha$  affected IL11-driven tumour growth (Fig. 3b and Extended Data Fig. 7). Furthermore, IL11 significantly promoted growth of 2/4 additional breast cancer cell lines despite low or undetectable levels of IL11R $\alpha$  (Fig. 3c, d).

Independence of tumour growth from direct stimulation of tumour cells by IL11 prompted us to investigate changes in the tumour microenvironment. IL11-driven tumours displayed higher intratumoral vascular density compared to parental ones (Fig. 3e, f), more dispersed patterns of collagen organization and had more stromal fibroblasts (Extended Data Fig. 8). Both increased vascularization and reorganization of the extracellular matrix have been implicated in the promotion of tumour growth<sup>21,22</sup>, suggesting that the tumour-promoting effects of IL11 may be attributable to microenvironmental changes.

### Clonal competition dynamics

Contexts of polyclonal tumours strongly inhibited the expansion of individual sub-clones in comparison to monoclonal tumours (Fig. 2c). This phenomenon is known as clonal interference: when multiple clones with higher than average fitness emerge in a population at the same time, they interfere with each other; this slows down the rate of clonal evolution<sup>16</sup>. However, the reduced expansion of individual sub-clones in IL11-driven polyclonal tumours could also be the result of a growing population. Therefore, to distinguish between the effects of clonal interference and expanding tumour volume, we determined clonal expansions in slower growing polyclonal tumours without IL11 (Fig. 2c). We found that while the removal of IL11 significantly affected clonal composition of the

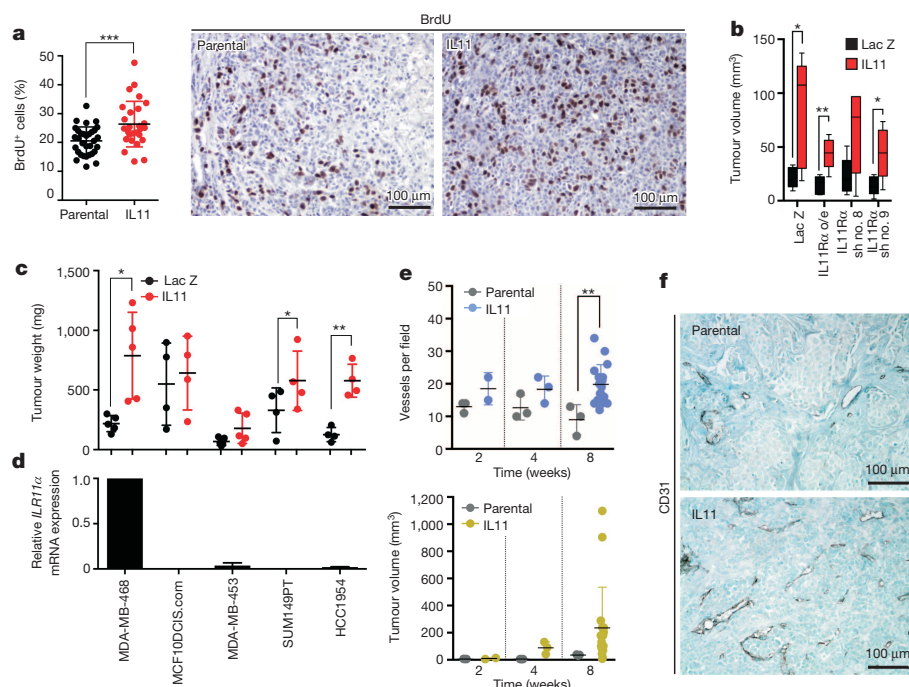
tumours ( $P < 0.0001$  for the interaction factor in a two-way ANOVA), expansion of most of the sub-clones remained inhibited. This indicates that clonal interference is a major determinant of the differences in the competitive dynamics in polyclonal tumours.

To investigate the rules of tumour growth and to predict clonal dynamics on a longer timescale, we then developed a mathematical framework incorporating clonal interference and heterogeneity. First, we investigated the growth behaviour of monoclonal tumours, finding that tumours exhibited an exponential growth pattern (Extended Data Fig. 3b). We then estimated the clone-specific exponential growth rates for each monoclonal growth experiment. With these rates we predicted tumour sizes in polyclonal tumours adding a dynamic interaction term (Fig. 4a, Extended Data Fig. 3c, d and Supplementary Information).

In order to account for interactions between a driver clone and other clones, we investigated a hierarchy of nested, increasingly complex mathematical descriptions of clonal dynamics for their ability to predict data from individual polyclonal growth experiments. The null hypothesis of no clonal interactions was easily rejected. The best agreement between model predictions and experimental observations in polyclonal tumours was achieved by including a constant positive growth effect of the IL11 clone on all other clones. Higher-order interactions involving multiple drivers did not improve the predictive power of the model. The best-fitting model was then used to predict heterogeneity in polyclonal tumours over longer timescales. In the absence of IL11, clonal heterogeneity was predicted to eventually vanish, as clones with the highest proliferation rates outcompete less fit clones. In contrast, non-cell-autonomous stimulation of cell growth supports clonal diversity over clinically relevant timescales (Fig. 4b).

As anti-cancer therapy exerts selective pressures that can affect evolutionary dynamics, we investigated the effect of treatment with doxorubicin, a commonly used chemotherapeutic agent in breast cancer, on the diversity of the tumour cell population. Two rounds of doxorubicin administration substantially inhibited tumour growth and cell proliferation in polyclonal tumours (Extended Data Fig. 6a, b). Instead of the expected changes in the expansion of specific sub-clones differing in drug sensitivity, we found that the amplitude of clonal expansion and



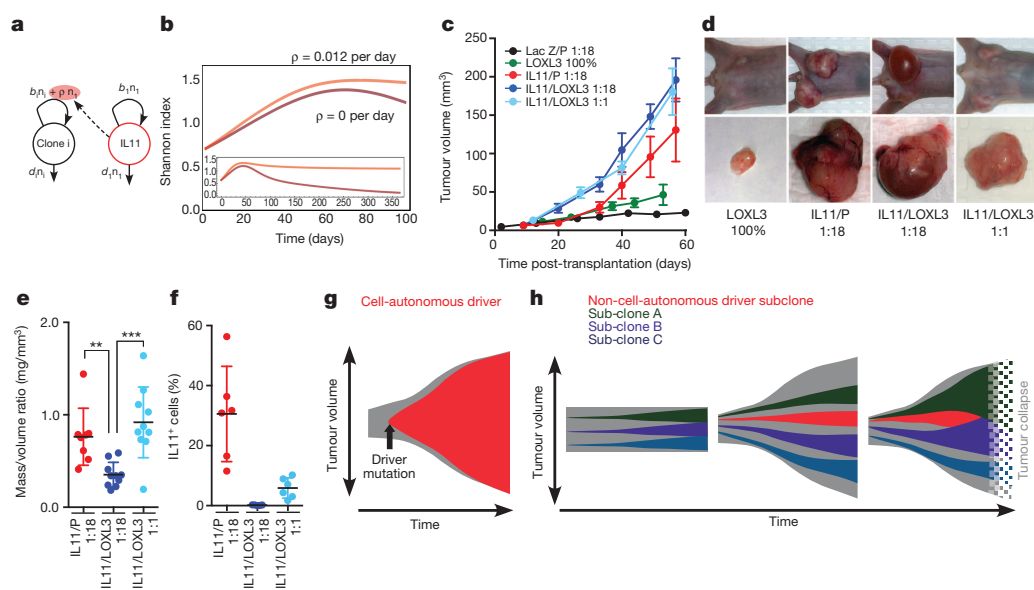


**Figure 3 | IL11 drives tumour cell proliferation via microenvironmental changes.** **a**, Quantification and representative images of anti-BrdU immunohistochemical staining in control and IL11-driven tumours. **b**, Tumour volumes 31 days post-transplantation of parental MDA-MB-468 cells, cells overexpressing or with downregulated IL11R $\alpha$ ,  $n = 5$  per group. **c**, Tumour weights of contralateral parental and IL11 expressing tumours formed by the indicated cell lines. **d**, Levels of expression of *IL11R $\alpha$*  mRNA in

indicated cell lines, normalized to MDA-MB-468. **e**, Quantification of average number of CD31 $^{+}$  vessels per field and tumour volumes. **f**, Representative images of anti-CD31 immunohistochemical staining. \* $P < 0.05$ , \*\* $P < 0.01$  and \*\*\* $P < 0.001$ , respectively of unpaired (**a**, **b**, **e**) or paired (**c**) Student's  $t$ -test. Error bars indicate s.e.m. Data shown are representative of at least 2 independent experiments.

contractions was increased compared to untreated tumours, reducing clonal diversity (Extended Data Fig. 6c, d). Therefore, even in the absence of selection for resistant subpopulations, doxorubicin treatment non-specifically amplified the effects of differences in competitive fitness. This observation was most probably a result of increased competition due to treatment-induced stabilization of the population size.

The lack of correlation between clonal expansion and tumour growth prompted examination of the competition between IL11 and LOXL3 sub-clones. The latter showed the strongest expansion in population frequency without being able to drive tumour growth (Fig. 2d). IL11 accelerated the growth of tumours with LOXL3 competitors beyond the growth rates seen with IL11/parental (IL11/P) controls (Fig. 4c),



**Figure 4 | Effect of IL11 on clonal dynamics.** **a**, Outline of the linear model that best explains polyclonal dynamics (see Supplementary Information). **b**, Prediction of diversity over time without (dark) or with (light) non-cell-autonomous driver. **c**, Tumour growth kinetics,  $n = 10$  per group. **d**, Representative images. **e**, Mass/volume ratios of tumours in **c–e** excluding

cyst fluid, each dot represents an individual tumour, \*\* $P < 0.01$ , \*\*\* $P < 0.001$ ; error bars indicate s.e.m. **f**, Final population frequencies of IL11 $^{+}$  cells in the indicated tumours. **g**, **h**, Models of cell-autonomous (**g**) and non-cell-autonomous (**h**) driving of tumour growth. Data shown are representative of at least 2 independent experiments.

consistent with the ability of faster proliferating LOXL3 cells to obtain additional benefit from IL11. However, upon sample collection, 1:18 IL11/LOXL3 tumours contained very little solid tissue. Most of the volume was filled with interstitial fluid, probably a remnant of necrotic liquefaction, whereas 1:18 IL11/P and 1:1 IL11/LOXL3 tumours remained solid (Fig. 4d, e).

Analysis of clonal composition revealed that LOXL3 had outcompeted the IL11 sub-clone below the detectability threshold in 1:18 IL11/LOXL3 tumours. In contrast, 1:1 IL11/LOXL3 tumours contained reduced, but substantial proportions of IL11 cells (Fig. 4f). Loss of IL11 cells most probably reflects differences in proliferation rates rather than apoptotic elimination of slower dividing cells seen in other experimental contexts<sup>23</sup>. We did not observe elevated rates of apoptosis in IL11 cells bordering LOXL3<sup>+</sup> cells in 1:1 IL11/LOXL3 tumours, and occasional IL11<sup>+</sup> cells could still be detected in 1:18 IL11/LOXL3 tumours (Extended Data Fig. 9). Additionally, the resulting clonal frequencies were consistent with predictions of our mathematical model (Supplementary Information). Most probably, elimination of IL11 sub-clone restored microenvironmental barriers, thereby prohibiting the maintenance of a large tumour. These findings provide experimental support for the idea that a clone responsible for driving tumour outgrowth can be outcompeted by a clone with faster proliferation, leading to tumour collapse<sup>24,25</sup>.

## Discussion

Widespread tumour heterogeneity challenges the common assumption that tumour growth and malignant phenotypes are driven by dominant clones that have the highest cell-autonomous fitness advantage (Fig. 4g). Previous studies in *Drosophila* and mouse models demonstrated that tumour growth can be supported by a small population of cells via direct non-cell-autonomous stimulation<sup>20,26,27</sup>. Furthermore, the cross-talk between sub-populations of tumour cells has been implied in metastasis<sup>28</sup>. Our results suggest that tumours can be driven by a sub-population of cells that does not have higher fitness, but instead stimulate growth of all tumour cells non-cell-autonomously by inducing tumour-promoting microenvironmental changes (Fig. 4h, middle). Conversely, non-cell-autonomous clonal expansion does not necessarily translate into increased tumour growth rates (Fig. 4h, left). The non-cell-autonomous driver sub-clone can be outcompeted by a sub-clone with higher proliferative output, thus collapsing the tumour (Fig. 4h, right). Notably, in our experiments IL11-expressing cells were initially intermingled with the competitors. Under the scenario of stochastic activation of expression, benefits of secretion of non-cell-autonomously acting factors might be skewed to the producer clone due to spatial considerations. Therefore, although extensive intermingling of evolutionarily diverged sub-populations has been reported for primary tumours<sup>29</sup>, it will be important to evaluate the effects of tumour topology in future studies.

Our results provide direct experimental evidence that clonal interference limits clonal expansions in tumours. Our modelling predicts that non-cell-autonomous driving of tumour growth can maintain clonal diversity over clinically relevant timeframes. In turn, clonal diversity can lead to clinically important phenotypic properties as suggested by the emergence of metastatic dissemination due to interactions between IL11- and FGF-expressing sub-populations. Non-cell-autonomous driving of tumour growth and inter-clonal interactions suggest that experimental analysis and clinical diagnostics focusing only on the most abundant sub-population of tumour cells might be misleading.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 27 November 2013; accepted 3 June 2014.

Published online 30 July 2014.

- Greaves, M. & Maley, C. C. Clonal evolution in cancer. *Nature* **481**, 306–313 (2012).
- Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).

- Fearon, E. R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* **61**, 759–767 (1990).
- Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
- Landau, D. A. *et al.* Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* **152**, 714–726 (2013).
- Ding, L. *et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**, 506–510 (2012).
- Burrell, R. A., McGranahan, N., Bartek, J. & Swanton, C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* **501**, 338–345 (2013).
- Anderson, K. *et al.* Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature* **469**, 356–361 (2011).
- Marusyk, A. & Polyak, K. Tumor heterogeneity: causes and consequences. *Biochim. Biophys. Acta* **1805**, 105–117 (2010).
- Merlo, L. M., Pepper, J. W., Reid, B. J. & Maley, C. C. Cancer as an evolutionary and ecological process. *Nature Rev. Cancer* **6**, 924–935 (2006).
- Weinberg, R. A. *The Biology of Cancer* (Garland Science, 2007).
- Bissell, M. J. & Hines, W. C. Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression. *Nature Med.* **17**, 320–329 (2011).
- Gatenby, R. A. & Gillies, R. J. A microenvironmental model of carcinogenesis. *Nature Rev. Cancer* **8**, 56–61 (2008).
- DeGregori, J. Challenging the axiom: does the occurrence of oncogenic mutations truly limit cancer development with age? *Oncogene* **32**, 1869–1875 (2013).
- Kreso, A. *et al.* Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. *Science* **339**, 543–548 (2013).
- Gerrish, P. J. & Lenski, R. E. The fate of competing beneficial mutations in an asexual population. *Genetica* **102–103**, 127–144 (1998).
- Putoczki, T. & Ernst, M. More than a sidekick: the IL-6 family cytokine IL-11 links inflammation to cancer. *J. Leukoc. Biol.* **88**, 1109–1117 (2010).
- Ernst, M. *et al.* STAT3 and STAT1 mediate IL-11-dependent and inflammation-associated gastric tumorigenesis in gp130 receptor mutant mice. *J. Clin. Invest.* **118**, 1727–1738 (2008).
- Putoczki, T. L. *et al.* Interleukin-11 is the dominant IL-6 family cytokine during gastrointestinal tumorigenesis and can be targeted therapeutically. *Cancer Cell* **24**, 257–271 (2013).
- Inda, M. M. *et al.* Tumor heterogeneity is an active process maintained by a mutant EGFR-induced cytokine circuit in glioblastoma. *Genes Dev.* **24**, 1731–1745 (2010).
- Bissell, M. J. & Radisky, D. Putting tumours in context. *Nature Rev. Cancer* **1**, 46–54 (2001).
- Folkman, J. Angiogenesis in cancer, vascular, rheumatoid and other disease. *Nature Med.* **1**, 27–31 (1995).
- Levayer, R. & Moreno, E. Mechanisms of cell competition: themes and variations. *J. Cell Biol.* **200**, 689–698 (2013).
- Maley, C. C., Reid, B. J. & Forrest, S. Cancer prevention strategies that address the evolutionary dynamics of neoplastic cells: simulating benign cell boosters and selection for chemosensitivity. *Cancer Epidemiol. Biomarkers Prev.* **13**, 1375–1384 (2004).
- Nagy, J. D. Competition and natural selection in a mathematical model of cancer. *Bull. Math. Biol.* **66**, 663–687 (2004).
- Wu, M., Pastor-Pareja, J. C. & Xu, T. Interaction between Ras<sup>V12</sup> and scribbled clones induces tumour growth and invasion. *Nature* **463**, 545–548 (2010).
- Cleary, A. S., Leonard, T. L., Gestl, S. A. & Gunther, E. J. Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature* **508**, 113–117 (2014).
- Calbo, J. *et al.* A functional role for tumor cell heterogeneity in a mouse model of small cell lung cancer. *Cancer Cell* **19**, 244–256 (2011).
- Navin, N. E. & Hicks, J. Tracing the tumor lineage. *Mol. Oncol.* **4**, 267–283 (2010).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank J. DeGregori, A. Goldman, A. Rozhok, M. Gonen and members of the Polyak and Michor laboratories for their critical reading of this manuscript and discussions. We thank L. Cameron in the DFCI Confocal Microscopy for her technical support. This work was supported by the Dana-Farber Cancer Institute Physical Sciences-Oncology Center (U54CA143798 to F.M.), CDRMP Breast Cancer Research Program W81XWH-09-1-0561 (A.M.), Cellex Foundation (V.A.), Deutsche Akademie der Naturforscher Leopoldina LPDS 2012-12 (P.M.A.) and the Breast Cancer Research Foundation (K.P.).

**Author Contributions** A.M. developed the experimental model, performed xenograft experiments and data analyses. D.P.T. performed immunohistochemical analyses and quantifications, and assisted with animal experiments. P.M.A. performed mathematical modelling and data analyses. V.A. assisted with image acquisition and analyses. K.P. supervised with help from F.M. All authors helped to design the study and write the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.P. ([kornelia.polyak@dfci.harvard.edu](mailto:kornelia.polyak@dfci.harvard.edu)).

## METHODS

**Cell lines.** Breast cancer cell lines were obtained from the following sources: MDA-MB-468, MDA-MB-453, and HCC1954 from ATCC; MCF10DCIS from Dr. F. Miller (Karmanos Cancer Institute, Detroit, MI), SUM149PT from Dr. S. Ethier, University of Michigan, Ann Arbor, MI), and 21NT from Dr. A. Pardee (Dana-Farber Cancer Institute, Boston, MA). Cells were cultured in media recommended by the provider, their identity confirmed by short tandem repeats (STR) analysis, and regularly tested for mycoplasma.

**Generation of MDA-MB-468 derivative lines ('sub-clones').** Entry cDNA ORFs in pDONOR223 or pENTR221 were obtained from human ORFeome collection v5.1 or Life Technologies, respectively. Lentiviral expression constructs were generated by Gateway swap into pLenti6.3/V5-Dest vector (Life Technologies) or pHAGE-EF (used for IL11 swap only, vector obtained from S. Elledge laboratory, Harvard Medical School) destination vectors and sequence verified. Assembling viral particles and transductions were performed following Life Technology protocols. Parental MDA-MB-468 cell lines were transduced with mCherry/Luciferase lentiviral construct (obtained from C. Mitsiades laboratory, DFCI) before derivation of specific sub-clones. Each derivative line was generated from a pool of  $1 \times 10^5$  to  $2 \times 10^5$  transduced cells. Lentiviral-mediated expression was verified by immunoblotting against V5 tag *in vitro* and further confirmed by immunohistochemistry *in vivo*. The GFP sub-clone was derived by lentiviral transduction of pLVX-AcGFP (Life Technologies).

**qPCR analysis of clonal composition.** The frequency of individual clones within tumours was determined by analysing the change in qPCR signal from the initial mixture, which was precisely defined through mixing of clones based on cell counts, and the terminal tumour. qPCR was performed using Life Cyclor 4800 (Roche) using SYBR green method with reaction mixtures purchased from Kapa Biosystems. Signals from individual clones were determined using a primer anchored in lentiviral backbone (anchor) and a primer specific for the clone-defining factor. As an internal reference we used primers specific for the peri-centromeric region of chromosome 12, which does not display copy number alterations in the MDA-MB-468 cell line. Primer sequences are listed below. The primers employed in the quantitation displayed linear amplification with >95% amplification efficiency. Change of frequency relative to the initial mixture was determined from Ct values for clone specific and internal reference qPCR signal based on ddCt method. Clonal proportions in polyclonal tumours were normalized based on total frequency of 1. For calculation of fold expansion, we used the clonality data to infer number of cells, following inferences between tumour mass and cellularity as described in the Supplementary Information.

**Target sequence of primers.** pLenti6.3/V5-Dest expressed: anchor TCCAGTGTGGTGAATCTCTG; IL11 CGTCAGCTGGGAATTGTGTC; SPP1 CATTCTGTGGGCTAGGAGA; VEGFC GAGCACTTGCCACTGGTGTGA; IHH GGTCTGATGTGGTGTGATGCC; HGF CTTTTCCTTTGTCCCTCTGTC; CCL5 CTGCTCC TCCAGATCTTTTGC; VEGFB CCATGAGCTCCACAGTCAAG; FIGF CTCCA CAGCTTCCAGTCTCTC; CXCL12 ATCTGAAGGGCAGAGTTTGG; VCAN GC GGAGAAATTCAGTGGTGT; SHH CCACATTGGGGATAAAGTGC; VEGFA GATTCTGCCCTCTCTCTTCT; CXCL14 TTTGGCTTCATTTCAGCTT; LOXL1 ACTATGAGCCCGAGTTGAGC; LOXL3 GTCTTCGATGTAGGCGTCTCT; AN GPTL4 GCGCCAGGACATTCATCT; IL6 GCGGCTACATCTTTGGAAATC; LA CZ CGGGCCTCTTCGCTATTAC; pLVX-AcGFP expressed; GFP F TCCTGGG CAATAAGATGGAG; GFP R TGGGGGTATTCTGCTGGTAG; pHAGE-EF-DEST expressed: anchor TGGGACGTCGTATGGGTATT; IL11 GGCTGCACCTGAC ACTTGAC; human-specific centromeric reference locus; F TTTGGGGCCTTAA CACTT; R AAGCAACCAAGCCCTTTCA.

**Xenograft experiments and doxorubicin treatment.** All animal procedures were approved by the DFCI ACUC (DFCI protocol#11-023) and followed NIH guidelines. Tumours were induced by bilateral orthotopic injection into 4–5-weeks old female *Foxn1*<sup>tm</sup> mice of  $1 \times 10^6$  cells resuspended in 50% Matrigel (BD Biosciences) per transplant. Animals without successful tumour grafting were excluded from the analysis. Tumour volumes were monitored by bi-weekly measurements of tumour diameters with electronic calipers. For doxorubicin treatment, animals were injected at days 15 and 22 post-transplantation with 5 mg per kg doxorubicin or PBS control. As tumour sizes distribution of control and treatment groups before treatment was similar, no randomization was performed. No blinding was performed during the tumour measurements in live animals.

**Immunoblot analysis.** A total of  $2 \times 10^6$  cells per sample were lysed in 100  $\mu$ l of RIPA buffer. 10  $\mu$ l of lysate was loaded per well of 4–12% Bis-Tris NuPage Midi gel (Life Technologies). Proteins were transferred to Immobilon PVDF membrane (EMD Millipore, Billerica, USA). Membranes were blocked for 30 min in StartingBlock blocking buffer (Thermo Scientific, Waltham, MA), then incubated overnight at 4°C with primary antibodies diluted 1:1,000 in PBST in presence of 2.5% BSA. After 3  $\times$  5 min washes, membranes were incubated with secondary antibodies at 1:20,000 dilution, washed 2  $\times$  5 min followed by a 20 min wash. The

membranes were developed with Immobilon substrate (EMD Millipore, Billerica, USA). The following antibodies were used:  $\beta$ -actin (Sigma, # A2228), IL11 $\alpha$  (R&D systems #MAB1977), HRP conjugated anti-mouse and anti-rabbit (Thermo Scientific).

**shRNA experiments.** shRNA constructs in pLKO lentiviral vectors were obtained from the Broad Institute RNAi consortium. shRNA with the following targeting sequences were used: IL11 $\alpha$  shRNA#4 CGGCAGATTCCACCTATAATT; IL11 $\alpha$  shRNA#5 TGGGACCATACCAAAGGAGAT; IL11 $\alpha$  shRNA#7 TGGAGCCA GTACCGGATTAAT; IL11 $\alpha$  shRNA#8 TGGCGTCTTTGGGAATCCTTT; IL11 $\alpha$  shRNA#9 ACTGATGAGGGCACCTACATC.

**IL11 ELISA.** Cells were plated at  $1 \times 10^5$  cells per well in a 6-well plate and left overnight at 37°C with 5% CO<sub>2</sub>. The next morning, the media was replaced and the cells returned to the incubator. After 5 h of incubation, the cells and the media were collected on ice in order to determine the concentrations of intracellular and secreted IL11, respectively. The harvested cells were counted, resuspended in PBS and lysed by rapid freeze thaw cycles. The media and cell lysates were used for human-IL11 ELISA (RayBiotech; ELH-IL11-001) according to the manufacturer's instructions. The values were adjusted for cell numbers as well as final volume to get an estimate of relative concentrations of IL11 in the two vector derivatives.

**Histological, immunohistochemical and multicolor immunofluorescence analyses.** For histological analyses, 5- $\mu$ m sections of formalin fixed paraffin embedded (FFPE) xenografts were stained with haematoxylin and eosin using standard protocols. For analyses of collagen content, the tumour sections were stained with Masson's trichrome stain kit (American Mastertech) following the manufacturer's instructions. Immunohistochemical analyses of bromodeoxyuridine (BrdU, Roche cat#11170376001, clone BMC9318, mouse monoclonal IgG<sub>1</sub>, 1:100), Ki-67 (Dako M724001, clone MIB-1, mouse monoclonal IgG<sub>1</sub>, 1:100), CD31 (Neomarkers RB10333, rabbit polyclonal, 1:50) and smooth muscle actin (SMA, Dako M085101, clone 1A4, mouse monoclonal IgG<sub>2a</sub>, 1:250) were performed using 5- $\mu$ m sections of FFPE xenografts. The tissues were deparaffinized and rehydrated. After heat-induced antigen retrieval in citrate buffer (pH 6 for BrdU and Ki-67) or Dako target retrieval solution (S2367, pH 9 for CD31 and SMA), the samples were blocked with 3% hydrogen peroxide in methanol followed by goat serum and stained with the primary for 1 h at room temperature. The samples were then incubated with anti-mouse or anti-rabbit IgG biotinylated antibody (1:100 dilution) for 30 min at room temperature followed by the ABC peroxidase system (Vectastain, ABC System Vector Laboratories). DAB (3,3'-diaminobenzidine) was used as the colorimetric substrate. The samples were washed twice with PBS-Tween 0.05% between incubations. Then the slides were counterstained with Harris haematoxylin or 1% methyl green. Scoring for the expression of each marker was done as follows: the percentage of Ki67<sup>+</sup> and BrdU<sup>+</sup> cells were estimated by counting an average of 1,500–2,000 cells per sample using ImageJ 1.45 s software from 4–6 randomly selected regions of the xenografts. Vessel density was scored by counting the number of CD31<sup>+</sup> vessels per 20 $\times$  field for 4–6 randomly selected fields in the tumour and the average was calculated. Blinding was used during key quantification analyses.

Multicolour immunofluorescence for cleaved caspase 3 (Cell Signaling cat#9661, rabbit monoclonal IgG, 1:50) and/or V5 (Invitrogen R960-25, mouse monoclonal IgG<sub>2a</sub>, 1:100) was performed similarly as above. After heat-induced antigen retrieval at pH 6, the samples were blocked with goat serum and stained with the primary overnight at 4°C followed by incubation with goat anti-rabbit IgG Alexa 488-conjugated (Life Technologies, 1:100 dilution, for detection of cleaved caspase 3) and goat anti-mouse IgG<sub>2a</sub> Alexa 555-conjugate (Life Technologies, 1:100 dilution, for detection of V5) for 45 min at room temperature. The samples were protected for long-term storage with VECTASHIELD HardSet Mounting Medium with DAPI (Vector laboratories, cat #H-1500). Before image analysis, the samples were stored at –20°C for at least 48 h. Different immunofluorescence images from multiple areas of each sample were acquired with a Nikon Ti microscope attached to a Yokogawa spinning-disk confocal unit using a 60 $\times$  plain apo objective, and OrcaER camera controlled by Andor iQ software. The montage images were created using the stitching plugin<sup>30</sup> in (Fiji Is Just) ImageJ 1.48f software.

**Terminal deoxynucleotidyl transferase dUTP nick end labelling (TUNEL) assay.** FFPE sections of the xenografts were deparaffinized and rehydrated. Sections were then treated with 60  $\mu$ g ml<sup>–1</sup> proteinase K (20 mg ml<sup>–1</sup>, Invitrogen, DNase- and RNase-free) in PBS for 15 min at room temperature. Protease digestion was stopped by consecutive washes in PBS and TdT buffer (Thermo Scientific). The sections were blocked with 3% hydrogen peroxide in methanol to inhibit endogenous peroxidase activity. TUNEL assays were performed at 37°C for 1 h in TdT buffer, 150 mM NaCl, 2  $\mu$ M biotin 16-dUTP (Roche) and 80 U per ml TdT (Thermo Scientific; EP0162). Following washing in PBS, labelled cells were visualized with the ABC peroxidase System (Vectastain, ABC System Vector Laboratories) using DAB (3,3'-diaminobenzidine) as the colorimetric substrate. The slides were counterstained with Harris haematoxylin. The percentage of TUNEL<sup>+</sup> cells were estimated

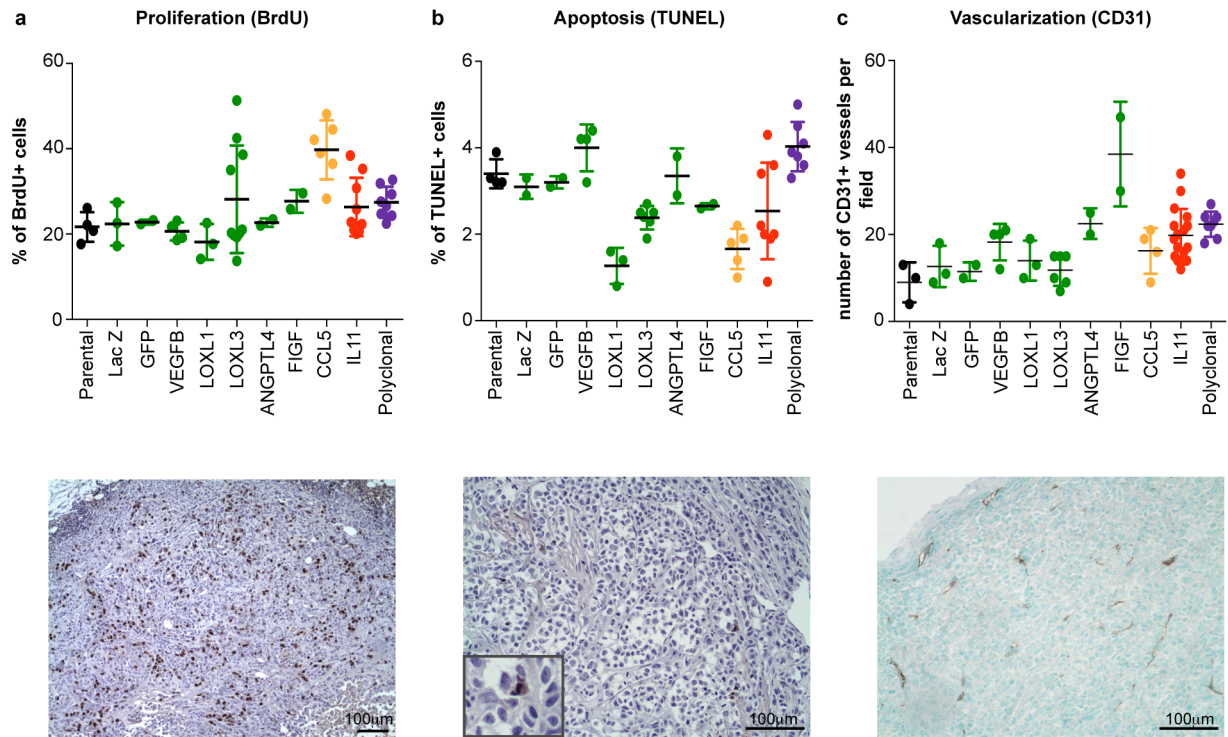


by counting an average of 600–1000 cells per sample using ImageJ 1.45 s software from 4–6 randomly selected regions of the tumours.

**Statistical analysis.** Sample size was determined based on pilot experiments followed by larger-scale studies to obtain significant differences (including the animal experiments). Estimation of variation within experimental group, normality test and statistical analyses indicated in figure legends were performed with Prism

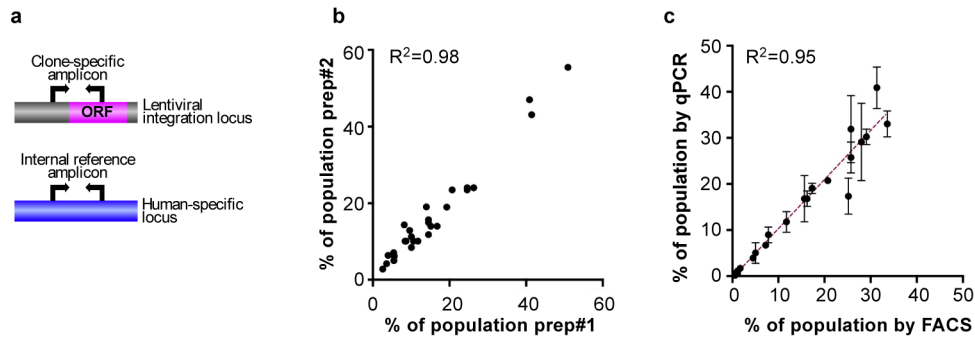
software (Graph Pad), or with Wolfram Mathematica. Unless otherwise specified, *P* values refer to the results of the two-tailed *t*-test.

30. Preibisch, S., Saalfeld, S. & Tomancak, P. Globally optimal stitching of tiled 3D microscopic image acquisitions. *Bioinformatics* **25**, 1463–1465 (2009).



**Extended Data Figure 1 | Proliferation, apoptosis and vascularization in selected groups.** a–c, Quantification and representative pictures of immunohistochemical analysis for markers of proliferation (a), apoptosis (b),

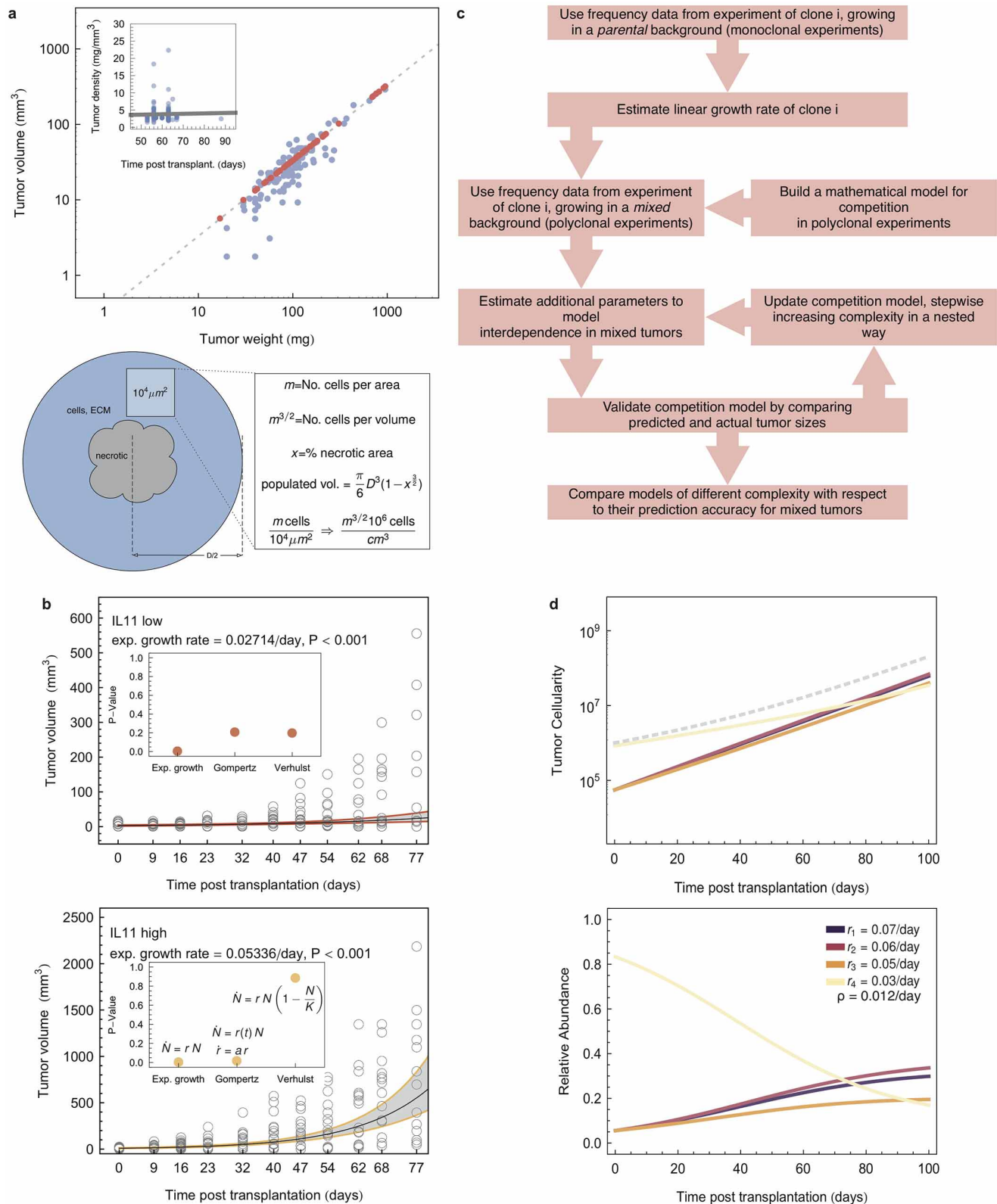
and vascularization (c). Each dot represents an individual tumour, error bars indicate s.d.



**Extended Data Figure 2 | Estimations of clonal frequencies.** **a**, Schematic outline of the quantification of clonal composition based on qPCR. Changes in clonal frequencies are determined based on changes in the ratios of clone-specific and a human-specific reference amplicon between initial mixtures and the resulting tumours. **b**, Reproducibility of clonality analysis between two different DNA preparations/qPCR from same tumour. **c**, Correlation between

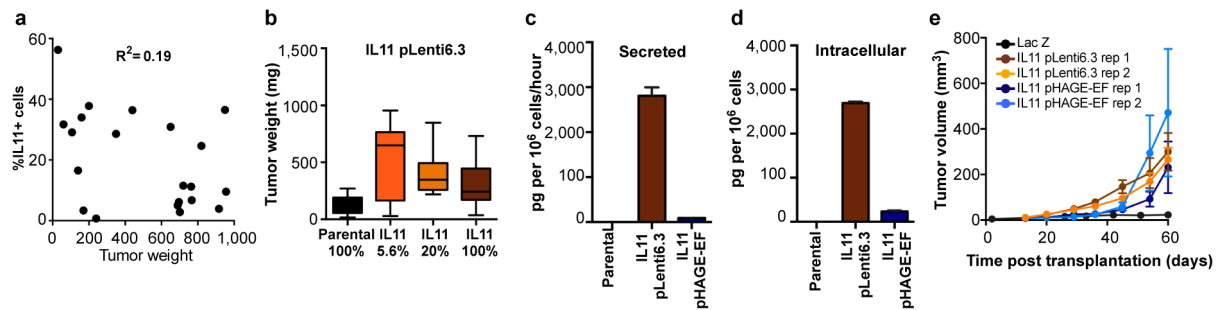
the results obtained using fluorescent-activated cell sorting (FACS) and qPCR based determination of clonal frequency after 6 weeks *in vitro* culture. Green fluorescent protein (GFP) labelled parental cells were mixed with individual sub-clones at initial ratios of 20:1.  $R^2$  indicates goodness of fit of linear regression.





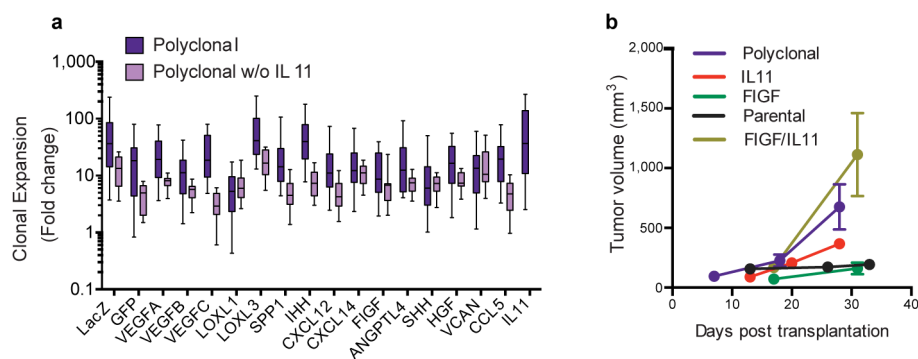
**Extended Data Figure 3 | Mathematical model.** **a**, Upper panel: estimation of tumour volume–density relation. The dashed line represents a linear regression with slope 0.33 ( $P < 0.01$ ). Red dots are predictions for which one value of the pair was missing. Inset, tumour density over time from clone-vs-parental competition experiments (dots). Tumour density did not correlate with the time of sample collection (line, linear regression with slope 0.012,  $P = 0.68$ ). Lower panel, schematic of estimation of cell numbers in tumour samples from two dimensional slices. **b**, Tumour volume over time from experiments (empty circles) and linear regression (exponential tumour growth law, black lines),

with 0.95 confidence intervals (grey areas). Inset: comparison of  $P$  values using different growth laws. **c**, Flow chart of mathematical modelling approach. **d**, Upper panel, growth dynamics under non-cell-autonomous driving, according to mathematical model (model B, see Supplementary Information), driver effect of IL11 was set to a typical value of 0.012/day. Example of four individual sub-clones (for example, IL11, LOXL3, slow-growing CCL5, LacZ), total tumour size indicated by dashed line; lower panel, frequency dynamics for the same set.



**Extended Data Figure 4 | Reproducibility and frequency-independence of tumour-growth promoting effects of IL11.** **a**, Relation between tumour weight and fraction of IL11 sub-clone cells upon tumour sample collection. **b**, Final weights of tumours initiated from the indicated mixtures of IL11 expressing and parental cells using pLenti6.3 backbone;  $n = 21$  for the 5.6% IL11,  $n = 10$  for the other groups. **c**, **d**, Secreted (pg per  $10^6$  cells per hour)

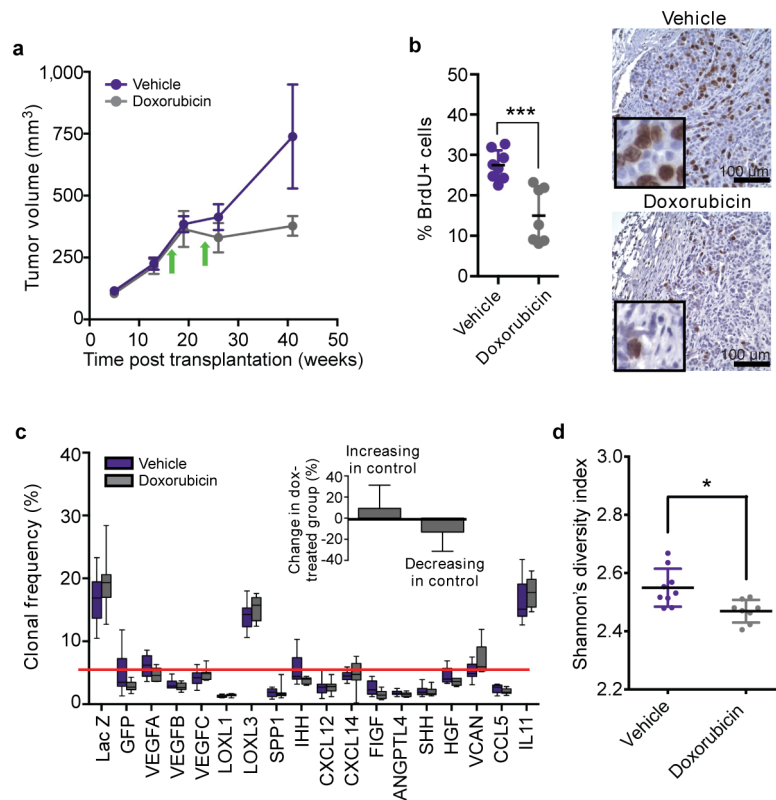
(c) and intracellular (pg per  $10^6$  cells) (d) levels of IL11 protein determined by ELISA in parental cells and in the IL11-expressing clones derived using the indicated lentiviral constructs. **e**, Growth kinetics of tumours initiated by transplantation of mixtures containing IL11-expressing cells from the indicated backbones competing with the parental cells.



**Extended Data Figure 5 | IL11 in clonal cooperation.** **a**, Expansion (fold-change over initial number of cells) of indicated sub-clones in the polyclonal tumours initiated with/without IL11 sub-clone,  $n = 10$  per group.

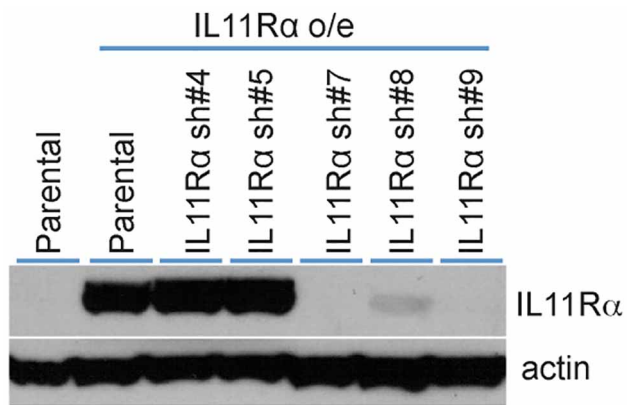
**b**, Growth curves of the tumours initiated by transplantation of the indicated groups, IL11 + FIGF indicates tumours initiated by 1:1 mixtures of IL11 and FIGF sub-clones.



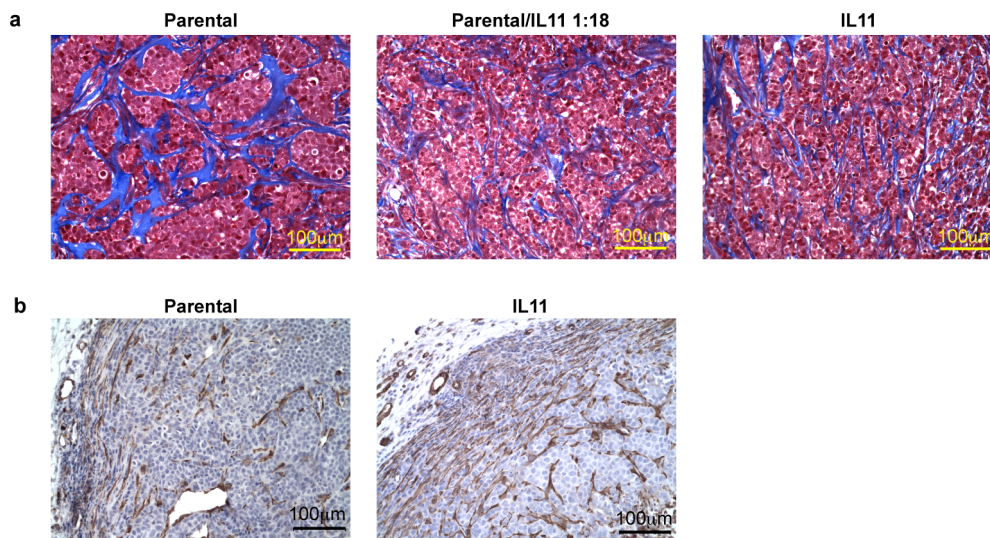


**Extended Data Figure 6 | The effects of doxorubicin on tumour growth and clonal composition.** **a–c**, Tumour growth (**a**), assessment of cell proliferation by BrdU staining (**b**) and clonal composition (**c**) of tumours initiated by polyclonal mixtures followed by treatment of the animals bearing established tumours with vehicle control or doxorubicin. Arrows mark intraperitoneal injections of doxorubicin (5 mg per kg) or vehicle. The inset in **c** quantifies

changes in frequency of clones expanding and shrinking compared to the initial frequencies. Interaction factor for two-way ANOVA between control and doxorubicin groups is statistically significant ( $P = 0.0059$ ). **d**, Shannon index for clonal diversity of vehicle and doxorubicin treated tumours,  $*P < 0.05$  in two-sample Kolmogorov–Smirnov test.



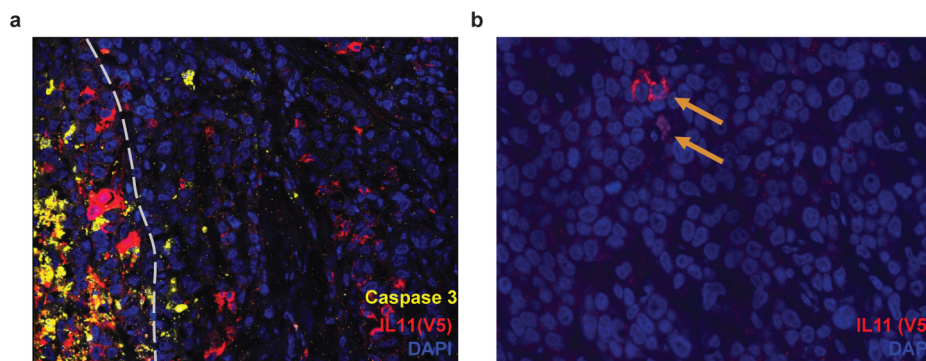
**Extended Data Figure 7 | Validation of IL11R $\alpha$  shRNA.** As the commercially available IL11R $\alpha$  antibodies are not sufficiently sensitive to detect endogenous IL11R $\alpha$  protein in the MDA-MB-468 cells, we tested the ability of shRNA to downregulate the expression of exogenously expressed IL11R $\alpha$ . Cells overexpressing IL11R $\alpha$  were stably transduced with IL11R $\alpha$ -targeting shRNAs and the expression of IL11R $\alpha$  and  $\beta$ -actin (loading control) were analysed by immunoblotting.



**Extended Data Figure 8 | The effects of IL11 on the tumour microenvironment.** **a**, Collagen organization in parental and IL11 expressing tumours. Representative images of collagen structure (blue) in the indicated

tumours as determined by tri-chrome staining. **b**, Smooth muscle actin positive (SMA) stromal cells in control and IL11 expressing tumours. Representative images of immunohistochemical staining for SMA.





**Extended Data Figure 9 | IL11 cells are not specifically eliminated in IL11/LOXL3 tumours.** **a**, Immunofluorescence analysis of apoptosis in 1:1 IL11/LOXL3 tumours. Apoptotic marker cleaved caspase 3 (yellow) indicates lack of increase in apoptosis in IL11 (red, V5<sup>+</sup>) cells bordering LOXL3 (V5<sup>-</sup>), as

LOXL3 cDNA has a stop codon before the tag). Grey dashed line demarcates the border of the necrotic area, where most of cell death occurs. **b**, Occasional IL11<sup>+</sup> cells (indicated by arrows) could still be detected in the remnants of 1:18 IL11/LOXL3 tumours.

**Extended Data Table 1 | List of factors employed in sub-clonal derivations**

Official gene symbol	Official gene name	Rationale for picking
LACZ	beta-D-galactosidase	Control
GFP	green fluorescent protein	Control
VEGFA	vascular endothelial growth factor A	Angiogenesis
VEGFB	vascular endothelial growth factor B	Lymphangiogenesis, metastasis
VEGFC	vascular endothelial growth factor C	Angiogenesis, lymphangiogenesis, and metastasis
LOXL1	lysyl oxidase-like 1	Invasion and metastasis
LOXL3	lysyl oxidase-like 3	Invasion and metastasis
SPP1	secreted phosphoprotein 1	Promotion of tumor growth through recruitment of bone marrow-derived cells
IHH	indian hedgehog	Activation of stroma
FIGF	c-fos induced growth factor	Lymphangiogenesis, metastasis
CXCL12	chemokine (C-X-C motif) ligand 12	Leukocyte infiltration, proliferation, metastasis
CXCL14	chemokine (C-X-C motif) ligand 14	Increased motility and invasiveness
SHH	sonic hedgehog	Promotion of tumor growth
VCAN	versican	Invasion, metastasis and growth
HGF	hepatocyte growth factor	Migration, adhesion and angiogenesis
CCL5	chemokine (C-C motif) ligand 5	Recruitment of monocytes
IL11	interleukin 11	Bone metastasis
ANGPTL4	angiopoietin-like 4	Angiogenesis and metastasis
IL6*	interleukin 6	Survival, proliferation

\* An IL6 expressing sub-clone was generated and tested in the pilot experiments but was excluded due to high systemic toxicity.

# Biogeography and individuality shape function in the human skin metagenome

Julia Oh<sup>1</sup>, Allyson L. Byrd<sup>1</sup>, Clay Deming<sup>1</sup>, Sean Conlan<sup>1</sup>, NISC Comparative Sequencing Program†, Heidi H. Kong<sup>2\*</sup> & Julia A. Segre<sup>1\*</sup>

**The varied topography of human skin offers a unique opportunity to study how the body's microenvironments influence the functional and taxonomic composition of microbial communities. Phylogenetic marker gene-based studies have identified many bacteria and fungi that colonize distinct skin niches. Here metagenomic analyses of diverse body sites in healthy humans demonstrate that local biogeography and strong individuality define the skin microbiome. We developed a relational analysis of bacterial, fungal and viral communities, which showed not only site specificity but also individual signatures. We further identified strain-level variation of dominant species as heterogeneous and multiphyletic. Reference-free analyses captured the uncharacterized metagenome through the development of a multi-kingdom gene catalogue, which was used to uncover genetic signatures of species lacking reference genomes. This work is foundational for human disease studies investigating inter-kingdom interactions, metabolic changes and strain tracking, and defines the dual influence of biogeography and individuality on microbial composition and function.**

Human skin harbours an abundant microbial ecosystem with bidirectional metabolic exchanges supporting symbiotic and commensal processes. The skin's surface consists of diverse microenvironments with distinct pH, temperature, moisture, sebum content and topography<sup>1</sup>. These niche-specific physiological differences influence the resident bacteria<sup>2,3</sup> and fungi<sup>4</sup>; oily surfaces like the forehead support lipophilic bacteria that differ from dry, low biomass sites like the forearm. In turn, microbial sensing and signalling mechanisms, metabolic pathways, or immunogenic features are likely to exhibit site-specificity to sustain host interactions. Similar to the distribution of skin microbes, skin disorders often present in a site-specific manner, such as atopic dermatitis (eczema) in arm and leg creases or psoriasis on the elbows and knees. Inter-kingdom and inter-species microbial interactions may exacerbate disease severity<sup>5</sup> or facilitate transitions from opportunistic to pathogenic. Although skin physiology is a dominant force, individuals retain unique elements of microbial profile and community organization. Here, we explore the complex skin microbial biogeography, integrating broad physiological characteristics with individual discriminatory attributes.

Studies based on phylogenetic marker genes (for example, bacterial 16S ribosomal RNA gene or fungal internal transcribed spacer (ITS) regions) have studied core taxonomic characteristics of different skin sites and disease states. However, such approaches survey kingdoms in isolation and provide limited information into an ecosystem's functionality. Metagenomic shotgun sequencing interrogates the full complement of DNA present in a sample, enabling characterization of both a community's functional capacity and genomes for which no targeted amplicon strategies exist. Several large-scale studies have used metagenomics to examine bacterial or viral communities of the healthy gut and other body sites<sup>6–8</sup>, or taxonomic and functional differences in type 2 diabetes<sup>9,10</sup>. To date, a systematic metagenomic investigation of human skin is lacking. The physiological heterogeneity and variable microbial biomass of the skin pose unique technical and analytical challenges for metagenomic studies. Each site on the human skin is constrained by ecological properties such as host microenvironment, yet possesses a distinct

biogeography that significantly influences microbial diversity, composition and biomass<sup>2–4,11</sup>.

We present the first systematic, multi-site metagenomic study of human skin. We determined the composition and function of the healthy skin microbiome using direct shotgun sequencing of 15 individuals at 18 clinically relevant sites, which included diverse skin microenvironments (dry, moist, sebaceous or toenail, Extended Data Fig. 1). Our dual approach incorporated reference-based and reference-free methods to characterize the metagenome. We present new insights into the larger community of skin microorganisms, including DNA viruses, lower eukaryotes, bacteria and subspecies of dominant bacteria. We defined how functional capacity varies by body site and created a multi-kingdom, skin-associated gene catalogue. Using new analytic approaches, we identified metagenomic 'clusters' representing species without known references. Our study demonstrates that biogeography and individuality significantly shape a community's functional and taxonomic characteristics and provides a framework for human studies investigating inter-kingdom interactions, metabolic changes and pathogen expansion in disease.

## Skin sampling and data characteristics

263 specimens were collected from 15 healthy adults (9 males, 6 females) from 18 defined anatomical skin sites (Supplementary Table 1). We modified previous clinical sample acquisition, DNA isolation and library preparation to generate shotgun metagenomic sequence data from skin sites, which varied in biomass and composition. For example, human-derived DNA accounted for  $19.4 \pm 6.7\%$  to  $98.2 \pm 0.1\%$  of reads, reflecting the difference between stratified, cornified plantar heel skin and nucleated inner nostril epithelium, respectively (Extended Data Fig. 2a). Microbial sequencing yields and estimated coverage also varied with skin physiological features ('microenvironment'), such that low-diversity, higher-biomass sebaceous sites generally achieving greater coverage (maximum  $81.0 \pm 7.0\%$ ) than high-diversity, lower-biomass dry or moist sites (minimum  $38.0 \pm 5.7\%$ , Extended Data Fig. 2c). We obtained a total of 289 gigabase pairs (Gbp) of non-human, quality filtered Illumina

<sup>1</sup>Translational and Functional Genomics Branch, National Human Genome Research Institute, NIH, Bethesda, Maryland 20892, USA. <sup>2</sup>Dermatology Branch, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, Maryland 20892, USA.

\*These authors contributed equally to this work.

†A list of authors and affiliations appears at the end of the paper.



microbial sequence reads (Extended Data Fig. 2a–c, Supplementary Table 1).

### Phylogenetic profiles of skin microbes

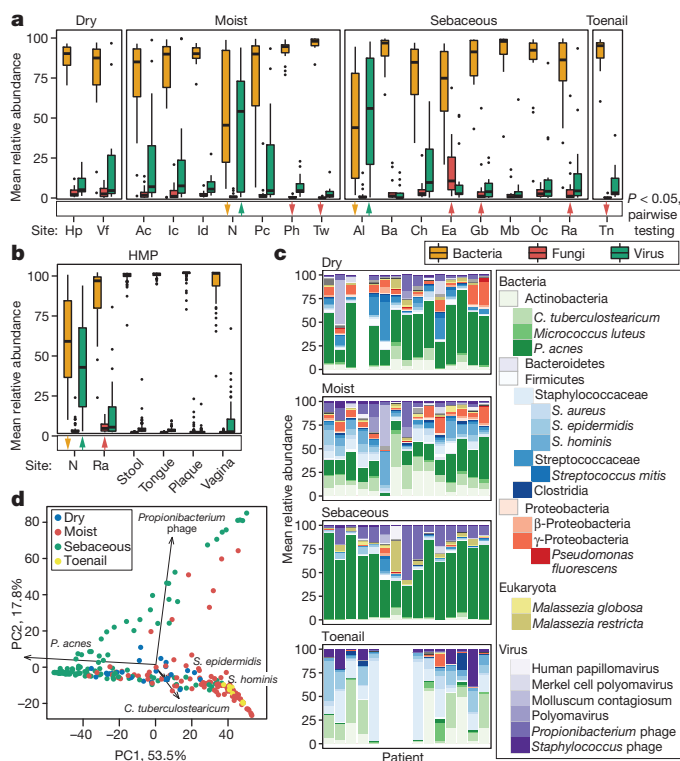
To explore the relative abundances of skin microbiota across kingdoms, we performed a relational analysis mapping reference reads to 2,342 bacterial, 389 fungal, 1,375 viral and 67 archaeal genomes. To validate taxonomic assignments, we compared our metagenomic data with 16S and ITS sequencing of the same samples, which showed high concordance (Extended Data Fig. 3, Supplementary Tables 3–5). While recognizing that fungal and viral genomes are more sparsely represented in reference databases, bacteria predominated at most sites (Fig. 1a–c, Extended Data Figs 1, 4a, Supplementary Table 6) and comprised the bulk of phylogenetic diversity with fungi and viruses contributing relatively fewer species. Fungi, primarily *Malassezia globosa* and *M. restricta*, were a lower fraction (3.9 ± 5.0%), except near the ears and forehead, which had a higher fungal presence (external auditory canal, 16.8 ± 5.1%; retroauricular crease 7.5 ± 4.2%; glabella 7.1 ± 4.0%). The feet had low fungal representation (plantar heel, 0.7 ± 0.2%; toenail 0.5 ± 0.3%; toe web 0.3 ± 0.1%), despite high diversity observed in amplicon-based studies. Archaea were nearly absent on skin, but DNA viruses were abundant at specific sites, with marked interpersonal variation. Note, RNA viruses are not interrogated by these methods and probably represent uncharacterized diversity. The nares and adjacent alar crease showed significant

viral representation (51.0 ± 11.8% and 54.6 ± 9.3%), compared to 9.9 ± 1.0% at other sites. Interestingly, a few individuals had sites that were dominated by viruses (up to 96%). These ‘blooms’ contained *Propionibacterium* or *Staphylococcus* bacteriophage and/or potential human viral pathogens (molluscum contagiosum, human papillomavirus, and Merkel cell polyomavirus), although skin sites were free of clinical lesions. Communities were shaped primarily by the microenvironment, in which differential abundance of stereotypical taxa such as *Propionibacterium acnes*, commensal staphylococci, *Corynebacterium* and *Propionibacterium* phage contributed most significantly to variation both between and within individuals (Fig. 1d).

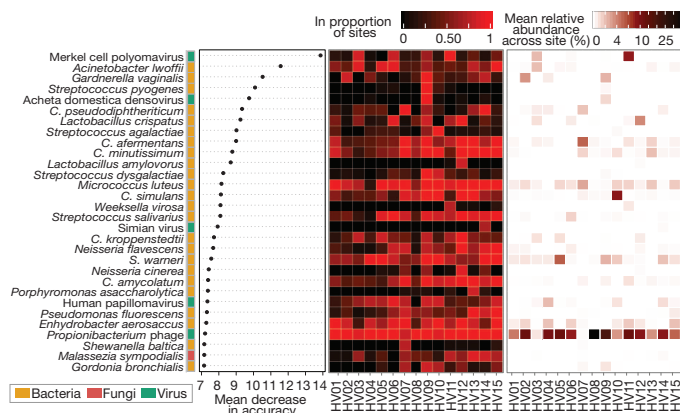
To compare skin with other body sites, we analysed 552 Human Microbiome Project (HMP) metagenomic samples obtained from the anterior nares, posterior fornix (vagina), retroauricular crease, stool, supragingival plaque and tongue dorsum (Fig. 1b, Extended Data Fig. 4b, Supplementary Tables 6, 7)<sup>12</sup>. Our skin samples were similar to those of the HMP in community membership and structure of all kingdoms ( $P > 0.05$ ). However, retroauricular crease samples from our study had greater fungal abundance than HMP (7.5% versus 3.4%), probably reflecting differences in nucleic acid extraction techniques, which we optimized to recover fungal DNA. Fungi were relatively scarce at non-skin sites. Similar to skin sites with phage co-occurring with their host bacteria, *Lactobacillus* phage was observed in the posterior fornix with marked interpersonal variation. Viruses were found in low abundance in the mouth, but *Streptococcus* phage was nearly universal, present in 99.2% of samples (mean abundance 1.2 ± 0.1%). Overall, the human body is rich in both bacterial and non-bacterial taxa, with site-specific fungal enrichment and viral blooms.

### Individuality underlies biogeography

Differential manifestations of phenotypes including disease susceptibility, antibiotic response, drug metabolism or even weight gain are likely to be influenced by an individual’s exclusive microbial community features. We explored whether we could classify individuals based on unique taxonomic signatures across their body. We used random forests, which incorporates interactions of both rare and abundant taxa, to identify key taxa that might differentiate individuals (Supplementary Table 8). Surprisingly, low-abundance taxa shared across skin sites discriminated individuals (Fig. 2). For example, the strongest discriminatory feature was Merkel cell polyomavirus, present in low abundance at all skin sites within one individual, regardless of site. Several taxa could also be discriminatory on an individual level; *Gardnerella vaginalis* and *Streptococcus pyogenes* were host-specific across all skin sites, in addition to taxa that probably represent transient populations (for example, *Acheta domesticus* densovirus).



**Figure 1 | Multi-kingdom relative abundances are strongly shaped by skin microenvironment.** **a**, Boxplots of mean relative abundance of different kingdoms by site. Black lines indicate median; boxes first and third quartiles. Triangles indicate significance (adjusted  $P < 0.05$ , Kruskal–Wallis post-hoc test) for over- (up) or under- (down) representation in a majority of pairwise comparisons between sites. Hp (hypothenar palm), Vf (volar forearm), Ac (antecubital crease), Ic (inguinal crease), Id (interdigital web space), N (nares), Pc (popliteal crease), Ph (plantar heel), Tw (toeweb space), Al (alar crease), Ba (back), Ch (cheek), Ea (external auditory canal), Gb (glabella), Mb (manubrium), Oc (occiput), Ra (retroauricular crease), Tn (toenail). **b**, Kingdoms in HMP body sites. **c**, Consensus relative abundance plots of major skin taxa by microenvironment. *C.*, *Corynebacterium*; *P.*, *Propionibacterium*; *S.*, *Staphylococcus*. **d**, Communities cluster primarily by microenvironment with sebaceous regions most distinct in principal components (PC) analysis.



**Figure 2 | Individual-specific signatures are typically low abundance but shared across most sites.** Left, variable importance plot of most discriminatory taxa from random forests analysis. For each individual, centre, proportion of the 18 sites in which each taxa is present, and right, mean relative abundance of that taxa across sites.

With our multi-kingdom taxonomy, we could differentiate our 15 individuals with >80% accuracy (19.3% error). The increased error estimates based upon kingdom-specific analyses (21.8%, bacteria; 74%, fungi; 41.2%, viruses) underscores the importance of understanding the full phylogenetic diversity of a community. Such approaches are relevant in identifying discriminatory features in disease states or assessing longitudinal community stability in which individuals may be identifiable by microbial features. While site-specificity serves as an overarching constraint on community composition, we observed a remarkable range of individual signatures within the skin biogeography.

### Strain heterogeneity in skin symbionts

We further explored individual signatures by examining strain-level variation; subspecies within a clade can possess different properties of transmissibility, virulence, antibiotic resistance, or metabolism<sup>13</sup>. To investigate strain-level heterogeneity, we focused on two common skin commensals with well-documented sequence variation, *P. acnes* and *S. epidermidis*. Using a reference-based approach that leveraged both single nucleotide polymorphisms and larger variants (Extended Data Fig. 5, Supplementary Tables 2, 9–12), we identified phylogenetically ‘most similar’ strains based on differentiating genomic features. To reduce false discovery, we characterized both strain and a more conservative subtype level that represents phylogenetically similar strain groups (Fig. 3a, b, Extended Data Figs 5, 6).

Given the extensive strain-level diversity observed for both species, our results suggest that individual and microenvironment differentially shape subspecies variation. *P. acnes* strains were more individual- than site-specific (Fig. 3c, e); 11/12 *P. acnes* subtypes were differentially abundant between individuals whereas only one differed between

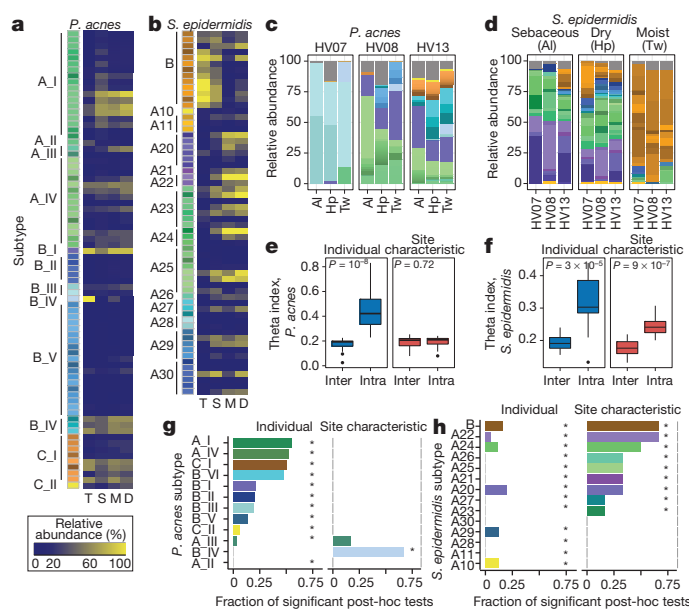
microenvironments (Fig. 3g). In contrast, *S. epidermidis* strains were significantly more site-driven with diminished inter-individual variation (Fig. 3d, f); nearly all subtypes were differentially abundant between sites (Fig. 3h) with subtype ‘B’ particularly dominant in the foot and toenail (Fig. 3b). These results strongly suggest that *P. acnes* and *S. epidermidis* communities are heterogeneous and multiphyletic, properties that probably vary by species and niche. Further analyses of this resolution will be powerful in determining genetic variation across time, topography and disease. In summary, our systematic analysis of microbial community composition has described a remarkable dynamism spanning inter-kingdom partnerships down to sub-species variability, characteristics that are driven both by broad ecological constraints and an individual’s unique carriage.

### Biogeography shapes functional diversity

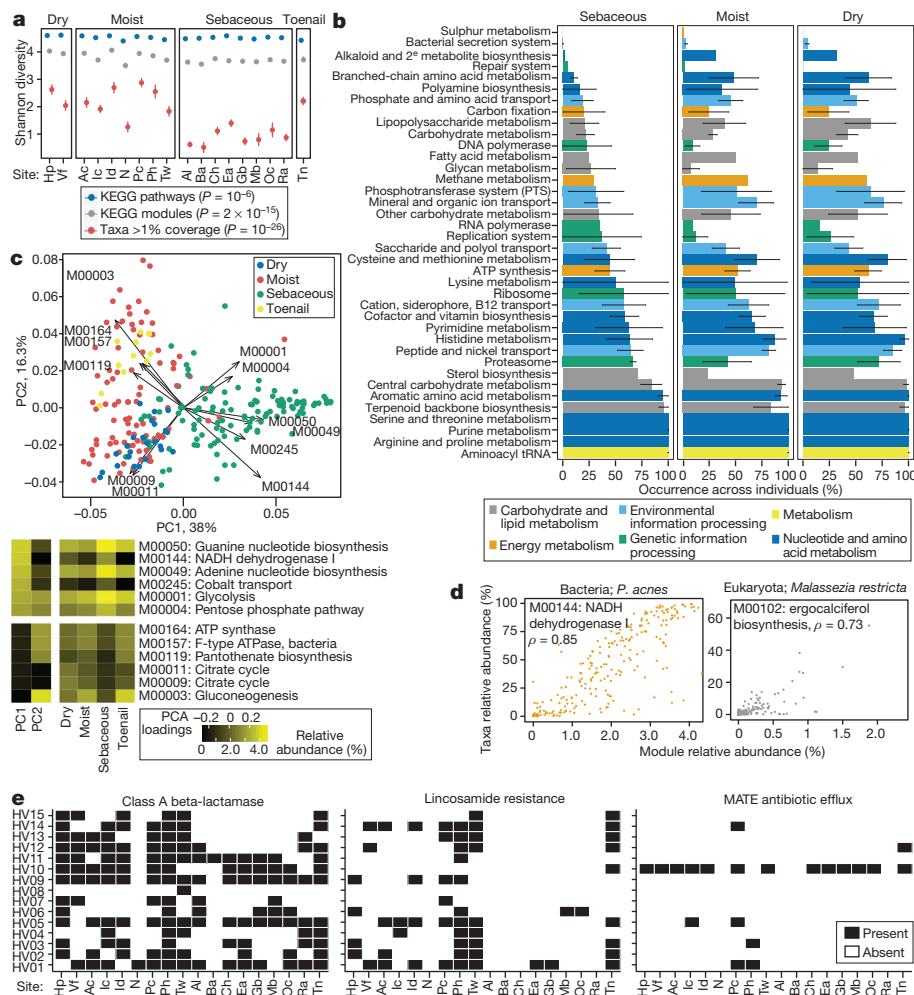
While taxonomy yields important insight into community organization, metagenomics also enables analysis of a community’s collective functional potential. Whereas previous studies reported that most metabolic pathways are evenly distributed across body sites<sup>12</sup>, we observed a modest decrease in metabolic diversity that occurred in tandem with lower taxonomic diversity in sebaceous sites (Fig. 4a). Investigating this concept of core functionality, we determined that only 30% (44/148) of modules were ‘core’ irrespective of site (present in  $\geq 2/3$  samples), representing processes essential to microbial growth and metabolism (Extended Data Fig. 7, Supplementary Tables 13–15). Extensive variability was observed within subclasses of major pathways, particularly transport systems (sulphate, glutamate, aspartame, L- or branched amino acids and sorbitol) and putrescine/spermidine biosynthesis and transport, which were typically absent in sebaceous regions, attesting to the chemical diversity likely to be present at higher-complexity sites. Conversely, most eukaryotic pathways were more prevalent in sebaceous sites (cell cycle, DNA replication, transcription, translation, protein degradation and vitamin D2 biosynthesis, a fungi-produced phytonutrient). Thus, although a strong functional core exists, this core metagenome can vary tremendously, reflecting functional diversification of skin microenvironments. Future studies with transcriptional profiling will probably reveal additional functional variance *in vivo*.

Modules present across all sites were typically low abundance and associated with uncharacterized biomolecular functions and metabolism. 88% of modules were differentially abundant in at least one microenvironment (adjusted  $P < 0.05$ , Supplementary Tables 13, 15), suggesting that functional capacity is driven primarily by biogeography. Principal components identified modules that discriminate microenvironments (Fig. 4c). Sebaceous sites (PC1) are distinguished by overrepresentation of glycolysis and related components (ATP and GTP generation) and NADH dehydrogenase I. Toenail samples differed primarily by the presence of different energy production components, such as conversion of oxaloacetate to fructose-6-phosphate, and ATPase and ATP synthase. Dry sites were characterized by the presence of citrate cycle modules. Covariance analysis imputing pathway abundance to select species suggested that *P. acnes* and *M. restricta* are likely candidates to drive some niche-specific metabolism, given their abundance in sebaceous sites (Fig. 4d, Extended Data Fig. 8).

With increasing concerns of antibiotic-resistant microorganisms, we explored the reservoir of antibiotic resistance genes in the skin. Although skin is physically compartmentalized from other body sites, cross-inoculation remains a risk factor. For example, the nares can harbour methicillin-resistant *Staphylococcus aureus* (MRSA)<sup>14</sup> underlying skin and soft tissue infections. Strain crosstalk between oral, lung and skin sites may underlie recurrent infections in immunocompromised patients<sup>15</sup>. Here, we identified presence/absence of well-characterized resistance gene families as pioneered for the gut<sup>16</sup> and soil<sup>17</sup>. We observed significant variability across individuals and resistance types (Extended Data Fig. 9, Supplementary Table 16). Certain antibiotic classes were highly host-specific, such as multi-antimicrobial extrusion (MATE) efflux pumps (Fig. 4e). In an example of site-specific dominance, lincosamide



**Figure 3 | *Propionibacterium acnes* and *Staphylococcus epidermidis* are heterogeneous and multiphyletic at the strain level.** **a, b**, Reference genomes used for *P. acnes* (**a**) and *S. epidermidis* (**b**). Leftmost bar shows subtypes (phylogenetically similar genomes) as colour groups. Adjacent heat map shows mean relative abundance by skin microenvironment. D, dry; M, moist; S, sebaceous; T, toenail. **c, d**, Select relative abundance plots; strain colours as in **a, b**. **e, f**, *P. acnes* subtypes differ more significantly between individuals than skin microenvironment with the converse observed for *S. epidermidis*. Boxplots of Yue–Clayton theta indices calculate similarity between (‘inter’) or within (‘intra’) individuals/microenvironments ( $\theta = 1$  means identical). Black lines indicate median, boxes show first and third quartiles.  $P$  value, Wilcoxon rank-sum test. **g, h**, Bar charts show *P. acnes* and *S. epidermidis* subtypes that differ by microenvironment or individual. Length of bar represents the fraction of post-hoc tests significant for each comparison; 105 comparisons for individual; 6 for microenvironment. \* $P < 0.05$ , adjusted Kruskal–Wallis test.



**Figure 4 | Functional capacity varies by microenvironment.** **a**, Shannon diversity of functional pathways and taxonomy by site;  $P$  value, Kruskal–Wallis test between microenvironments. Error bars, standard error of the mean. **b**, Microenvironments possess different core modules; ‘core’ means occurrence in more than 2/3 of samples. Error bars show variation within a class of modules (full version in Extended Data) that may arise from a unique specialization for that microenvironment. **c**, PCA shows clustering by microenvironment,

with strong separation of sebaceous, dry and toenail modules. Heat maps: left, loadings for the first two PCs; right, mean relative abundances for modules with the greatest variation by microenvironment. **d**, A module’s taxonomic origin can be imputed by Spearman correlation ( $\rho$ ; adjusted  $P \leq 2 \times 10^{-16}$ ) with *P. acnes* and *M. restricta* relative abundances. **e**, Presence of select antibiotic resistance gene families by individual and site.

resistance showed significant representation in three foot sites but was generally absent in sebaceous regions. Finally, certain families were broadly represented across samples, such as class A beta-lactamases, rRNA methyltransferases, efflux mechanisms, or quinolone resistance. Thus, carriage of antibiotic resistance families demonstrated both site- and individual-specificity, although we note that resistance activity may differ *in vivo*.

### Insights into microbial dark matter

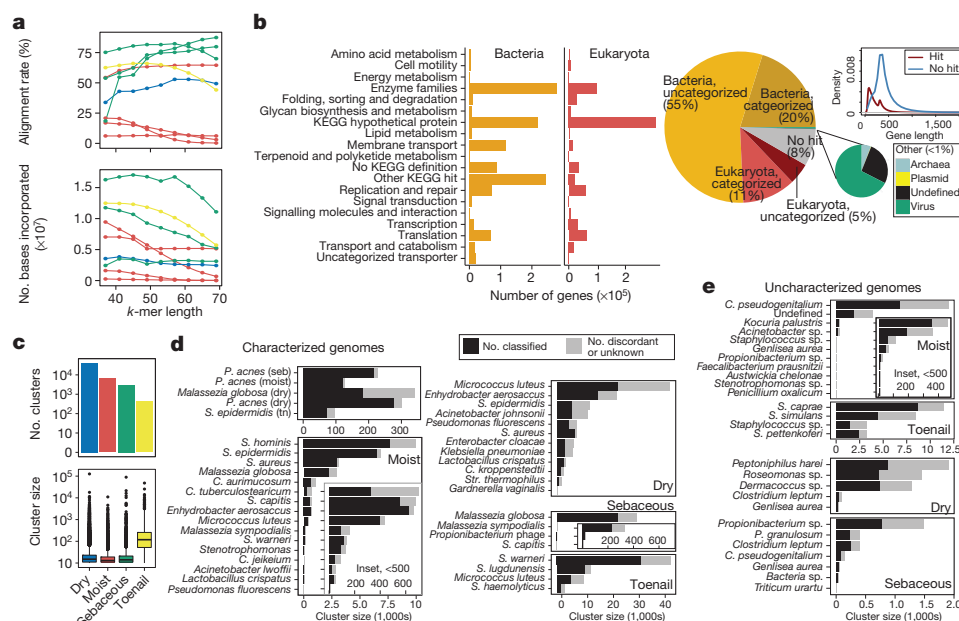
Our reference-based analysis showed a large variable fraction of reads (2–96%) unmapped to reference genomes, most frequently originating from decreased bacterial assignments (Supplementary Table 6, Extended Data Fig. 10a). Such uncharacterized sequences likely originate from both taxa with no representative reference and intraspecies pangenomic variation, which can represent significant gene content<sup>13</sup>. Using reference-free methods to capture this ‘dark matter’ of the skin metagenome, we created a skin gene catalogue that we then used to identify previously uncharacterized taxa in the skin. Such resources will be invaluable for downstream analyses, enabling *in silico* prediction and synthesis of genes and pathways that are over- or underrepresented in, for example, disease states.

The inherent variation in skin community complexity and human DNA admixture presents new challenges in reference-free methodologies;

variable microbial load and taxonomic diversity across sites affect sequencing depth and coverage. To account for this variability, we devised an adaptive and iterative strategy (Extended Data Fig. 10b, c) that optimizes assembly on a per-sample basis (Fig. 5a, Supplementary Table 17). We then established the first multi-kingdom skin microbial gene catalogue using both fungal and bacterial prediction models. Of 5.92 million open reading frames (ORFs), 75.7% could be reconstructed as bacterial and 15.9% as eukaryotic, consistent with our taxonomic analyses (Fig. 5b, Supplementary Table 18). Large numbers of KEGG (Kyoto Encyclopedia of Genes and Genomes) hypothetical genes (25.7% of bacterial, 48.3% of eukaryotic) are likely to represent pangenomic loci of characterized taxonomies, for example, *P. acnes* and *M. globosa*, based on association without pathway annotation. In support of their authenticity, ORFs with no identifiable homologues (7.9%) were typically longer than classified ORFs (Fig. 5b, inset). Less than 1% of ORFs were assigned to Archaea and viruses (which require unique prediction models), possibly reflecting integrative viruses or overlap in gene prediction models.

Finally, we used our gene catalogue to identify microbial species and pangenomic content independently of reference genomes. Under the assumption that genes from one genome covary in abundance across samples owing to physical linkage, we created metagenomic ‘clusters’<sup>9,10</sup> by correlating gene abundances across samples (Supplementary Table 18).





**Figure 5 | Reconstruction of metagenomic dark matter with reference-free methods.** **a**, Per-sample iterative assembly with variable  $k$ -mers (nucleotide words of length  $k$ ) optimizes assembly quality as assessed by metrics such as % reads mapping back to assembly (left) and the number of bases incorporated (right). Colours are as in **c**. **b**, Skin gene catalogue was mapped to the NCBI non-redundant (nr) database and KEGG to identify kingdom and functional category. Density plot compares length of genes with and without homology; gene length was typically larger for unmapped genes.

Most resultant clusters were relatively small, but others contained hundreds of thousands of predicted ORFs, which probably represent both genes and gene fragments. High-complexity dry sites had the most clusters and whereas toenails had the fewest, their median gene recruitment was significantly larger (Fig. 5c). To strengthen the reliability of our metagenomic clusters, we required clusters to share  $>50\%$  consensus taxonomy at the species level and uncovered large clusters of fungi, bacteria and viruses (Fig. 5d). *M. globosa*, *P. acnes* and *S. epidermidis* had very large clusters, consistent with their high abundance in skin. In addition to clusters representing referenced genomes, we also identified multiple uncharacterized genomes (Fig. 5e), most commonly species of common genera in the skin, including *Corynebacterium*, *Propionibacterium* and *Staphylococcus*. In summary, leveraging reference-free approaches, we identified previously undefined elements of the human skin microbiota. While dominant species or pathogens are targeted for sequencing, metagenomic studies reveal remarkable additional taxonomic and thereby functional diversity.

## Conclusions

The healthy skin metagenome possesses surprising taxonomic and functional diversity dependent on both biogeography and individuality. In contrast to other body sites like the gut, the skin has markedly higher viral and fungal representation. For most individuals, common skin species exist as a heterogeneous mix of strains, raising questions of whether transitions to a pathogenic state are mono- or multiphyletic, and how strain heterogeneity affects disease incidence or severity. Significant decreases in community diversity are a hallmark of a disease state<sup>18</sup>; whether such shifts occur at all taxonomic levels down to the subspecies awaits investigation. Our reference-based toolkit for multi-kingdom analyses and strain differentiation is broadly applicable to ecosystems with a well-characterized sequence space. Our reference-free resources, generated by adaptive assemblies, enable interrogation of the significant uncharacterized proportion of the metagenome, even identifying species without reference genomes.

**c**, Metagenomic clusters represent genes that covary in abundance across samples within a microenvironment; boxplots show cluster sizes; histograms show number of clusters ( $\log_{10}$  scale). **d**, A lowest common ancestor (LCA) was assigned to a cluster with  $>50\%$  consensus taxonomy. Bar length indicates the total number of 'genes' in a cluster; black represents the number of genes mapping to the LCA. Grey represents ambiguous or unannotated genes. 'Characterized' indicates that a reference genome exists for that species; for **e**, 'Uncharacterized genomes', no reference exists. Seb, sebaceous; tn, toenail.

From a therapeutic perspective, the metagenome represents a rich resource for synthetic biology approaches to modify and transplant endogenous elements to other communities. Studies of metabolic capacity, pathogenicity islands and virulence genes in disease states, with our catalogue from healthy skin, will uncover biomarkers associated with transmission, recurrence and severity of disease. Finally, characterization and tracking of surprisingly pervasive antibiotic resistance elements will remain clinically relevant, as skin sites can serve as a taxonomic and genetic reservoir for pathogens. We envision a new therapeutic landscape leveraging unique metagenomic profiles with tailored clinical interventions that reshape our microbial communities.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 3 June; accepted 21 August 2014.

- Grice, E. A. & Segre, J. A. The skin microbiome. *Nature Rev. Microbiol.* **9**, 244–253 (2011).
- Grice, E. A. *et al.* Topographical and temporal diversity of the human skin microbiome. *Science* **324**, 1190–1192 (2009).
- Costello, E. K. *et al.* Bacterial community variation in human body habitats across space and time. *Science* **326**, 1694–1697 (2009).
- Findley, K. *et al.* Topographic diversity of fungal and bacterial communities in human skin. *Nature* **498**, 367–370 (2013).
- Peters, B. M. & Noverr, M. C. *Candida albicans*-*Staphylococcus aureus* polymicrobial peritonitis modulates host innate immunity. *Infect. Immun.* **81**, 2178–2189 (2013).
- Arumugam, M. *et al.* Enterotypes of the human gut microbiome. *Nature* **473**, 174–180 (2011).
- De Vlaminc, I. *et al.* Temporal response of the human virome to immunosuppression and antiviral therapy. *Cell* **155**, 1178–1187 (2013).
- Handley, S. A. *et al.* Pathogenic simian immunodeficiency virus infection is associated with expansion of the enteric virome. *Cell* **151**, 253–266 (2012).
- Qin, J. *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55–60 (2012).
- Karlsson, F. H. *et al.* Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* **498**, 99–103 (2013).
- Grice, E. A. *et al.* A diversity profile of the human skin microbiota. *Genome Res.* (2008).

12. The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).
13. Tettelin, H. *et al.* Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial 'pan-genome'. *Proc. Natl Acad. Sci. USA* **102**, 13950–13955 (2005).
14. von Eiff, C., Becker, K., Machka, K., Stammer, H. & Peters, G. Nasal carriage as a source of *Staphylococcus aureus* bacteremia. *N. Engl. J. Med.* **344**, 11–16 (2001).
15. Oh, J. *et al.* The altered landscape of the human skin microbiome in patients with primary immunodeficiencies. *Genome Res.* (2013).
16. Sommer, M. O. A., Dantas, G. & Church, G. M. Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science* **325**, 1128–1131 (2009).
17. Forsberg, K. J. *et al.* The shared antibiotic resistome of soil bacteria and human pathogens. *Science* **337**, 1107–1111 (2012).
18. Kong, H. H. *et al.* Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. *Genome Res.* **22**, 850–859 (2012).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank D. Schoenfeld, A. Pradhan, M. Park and G. Bouffard for their efforts. We also thank members of the Segre laboratory and M. C. Udey for their discussions. This work was supported by National Institutes of Health (NIH) NHGRI and NCI Intramural Research Programs and in part by 1K99AR059222 (H.H.K.). This study used the high-performance computational capabilities of the NIH Biowulf Linux cluster. Sequencing was funded by grants from the National Institutes of Health (1UH2AR057504-01 and 4UH3AR057504-02).

**Author Contributions** J.O., H.H.K. and J.A.S. designed the study. H.H.K. collected patient samples. C.D. prepared the clinical samples for sequencing, which was carried out by

the members of the NIH Intramural Sequencing Center Comparative Sequencing program. J.O., A.L.B. and S.C. analysed sequence data. J.O., H.H.K. and J.A.S. drafted the manuscript. All authors read and approved the final version of the manuscript.

**Author Information** Data deposition is with the SRA and all sequences can be accessed under BioProject 46333. Human subject clinical data are deposited with dbGaP phs000266. Analysis workflow is available at <https://github.com/julia0h/skinmetagenome.git>. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.H.K. ([konghe@mail.nih.gov](mailto:konghe@mail.nih.gov)) or to J.A.S. ([jsegre@mail.nih.gov](mailto:jsegre@mail.nih.gov)).

#### NISC Comparative Sequencing Program

Betty Barnabas<sup>1</sup>, Robert Blakesley<sup>1</sup>, Gerry Bouffard<sup>1</sup>, Shelise Brooks<sup>1</sup>, Holly Coleman<sup>1</sup>, Mila Dekhtyar<sup>1</sup>, Michael Gregory<sup>1</sup>, Xiaobin Guan<sup>1</sup>, Jyoti Gupta<sup>1</sup>, Joel Han<sup>1</sup>, Shi-ling Ho<sup>1</sup>, Richelle Legaspi<sup>1</sup>, Quino Maduro<sup>1</sup>, Cathy Masiello<sup>1</sup>, Baishali Maskeri<sup>1</sup>, Jenny McDowell<sup>1</sup>, Casandra Montemayor<sup>1</sup>, James Mullikin<sup>1</sup>, Morgan Park<sup>1</sup>, Nancy Riebow<sup>1</sup>, Karen Schandler<sup>1</sup>, Brian Schmidt<sup>1</sup>, Christina Sison<sup>1</sup>, Mal Stantripop<sup>1</sup>, James Thomas<sup>1</sup>, Pamela Thomas<sup>1</sup>, Meg Vemulapalli<sup>1</sup> & Alice Young<sup>1</sup>

<sup>1</sup>NIH Intramural Sequencing Center, National Human Genome Research Institute, Bethesda, Maryland 20852, USA.

## METHODS

**Subject recruitment and sampling.** Healthy male and female volunteers of 23 to 39 years of age without chronic skin diseases were recruited from the Washington DC metropolitan region, USA, between June 2011 and May 2013. This natural history study was approved by the Institutional Review Board of the National Human Genome Research Institute (<http://www.clinicaltrials.gov/ct2/show/NCT00605878>). All subjects provided written informed consent before participation. Subjects provided medical and medication history and underwent a physical examination. Exclusion criteria included history of chronic medical conditions, including chronic dermatologic diseases, and use of antimicrobial medication (antibiotic or antifungal treatments) 1 year before sampling. Cleansing with only non-antibacterial cleansers was allowed during the 7 days before sample collection. To maximize microbial load, no bathing, shampooing or moisturizing was permitted within 24 h of sample collection<sup>15</sup>, which we have previously observed produces no discernible shifts in the overall diversity and structures of skin communities.

18 skin sites representing diverse physiological characteristics and sites of predilection for specific dermatologic diseases were sampled: moist (antecubital crease, inguinal crease, interdigital web space, nares, popliteal crease, plantar heel, toe web space), dry (hypothenar palm, volar forearm), sebaceous (alar crease, back, cheek, external auditory canal, glabella, manubrium, occiput, retroauricular crease), and toenail (Extended Data Fig. 1). Additional unmatched samples excluded from statistical analyses included samples extracted with the NEBNext Microbiome DNA Enrichment Kit (NEB), axillary vault (moist), bacterial and fungal mock communities<sup>19</sup>, samples that were whole-genome-amplified before library creation, and samples from disease patients (STAT3-hyper IgE, SH). To obtain sufficient DNA from defined anatomical skin sites with low and variable microbial biomass, we modified clinical sample acquisition methods using a swab-scrub-swab procedure, in which a defined anatomical skin area was swabbed with a swab (Catch-All Sample Collection Swabs, Epicentre) pre-moistened with yeast cell lysis buffer (MasterPure Yeast DNA Purification Kit, Epicentre), scraped via sterile disposable surgical blade, and swabbed with the same swab again. Residuals from the scalpel and swab were collected into lysis buffer. Nares and external auditory canal sites were sampled via swabbing with pre-moistened swabs that were then placed into lysis buffer. Toenail samples were cut with sterilized nail clippers and placed into lysis buffer. All samples were stored at  $-80^{\circ}\text{C}$  until extraction. Samples were then incubated in yeast cell lysis buffer (MasterPure Yeast DNA Purification Kit, Epicentre) and treated with Readylyse (Epicentre) for 30 min at  $37^{\circ}\text{C}$ , then mechanically disrupted using 5 mm stainless steel beads (Qiagen) in a TissueLyser (Qiagen) for 2 min, 30 Hz. Samples were incubated for 30 min at  $65^{\circ}\text{C}$ , placed on ice for 5 min, and debris spun down after treatment with MPC protein precipitation reagent. Samples were combined with 350  $\mu\text{l}$  of 100% ethanol and column purified using the Invitrogen PureLink Genomic DNA. Finally, samples were eluted in 30  $\mu\text{l}$  of water (MoBio).

**Sample sequencing.** Because of low bioburden typical of skin samples, Illumina libraries were created using Nextera library preparation. Briefly, 1–50 ng of extracted DNA was used as input into the transposome fragmentation step. Manufacturer's protocol was followed with the exception of using 10 cycles of PCR. 1–10 ng of extracted DNA was used as input according to manufacturers' recommended protocol (Qiagen Repli-G Mini). Libraries were then sequenced with  $2 \times 100$  bp paired end reads on an Illumina HiSeq at the NIH Intramural Sequencing Center with a target of 15 or 50 million clusters, depending on the microbial diversity of that site and the human DNA admixture. To ascertain that the Nextera approach resulted in minimal sequencing bias, we calculated expected distribution of breaks as represented by the expected frequency of pentamers starting a read for four different genomes, with high correlation with a standard Illumina prep. Moreover, expected versus observed frequencies of species in sequencing of the bacterial mock community were closely matched.

In total, we obtained 7.4 billion reads (289 Gbp) of non-human, quality-filtered paired-end and singleton reads (median 9.5 million reads (893 Mbp) per sample, mean insert size  $145 \pm 2$  bp). Sequencing data were processed to remove low quality reads and any read pairs in which at least one read mapped to the human hg19 human reference. Nextera adaptor sequences were trimmed, if necessary, using Crossmatch 1.090518 (<http://www.phrap.org>) and custom scripts. Bases with quality score below 20 were trimmed, and reads  $< 50$  bp length were removed. Sequencing depth varied by site with estimated  $k$ -mer coverage ranging from  $38.0 \pm 5.7\%$  to  $81.0 \pm 7.0\%$  based on the accumulation of unique DNA substrings, or  $k$ -mers. Rarefaction curves were generated using Khmer v0.7.1<sup>20</sup> with a  $20\times$  coverage cut-off. Briefly, reads were split into  $k$ -mers, compared to a  $k$ -mer coverage table and kept only if the median  $k$ -mer coverage was below the cutoff. Resulting curves showed the coverage of  $k$ -mer space as a function of sequencing effort. Median insert size was estimated from a subsample of paired reads that match hg19. Post sequence quality control, samples with  $> 20$  million reads remaining were subsampled to 10 million paired end reads, and singletons were discarded. HMP data from the anterior nares, retroauricular crease, stool, posterior fornix, tongue dorsum and

supragingival plaque were obtained from <ftp://public-ftp.hmpdacc.org> and subsampled to 1 million reads for taxonomic comparisons.

**Amplicon processing.** To validate our taxonomic assignments, normalize for sequencing levels, and reduce false positives, we also compared our results with matched bacterial 16S and fungal ITS amplicon sequencing. 159 matched 16S rRNA and 92 matched ITS1 samples were processed as previously described<sup>15</sup>. Briefly, the V1–V3 region of the 16S rRNA gene was amplified using the barcoded 27F and 534R and the ITS1 with 18SF and 5.8S–1R primers. Amplicon libraries were sequenced on a 454 GS FLX (Roche) instrument using titanium chemistry. 16S rRNA and ITS1 samples were processed using the mothur pipeline<sup>21</sup> as previously described<sup>15</sup>. Briefly, 454 flow gram data were denoised, error-trimmed, and chimaeric sequences removed. 16S sequences were classified using RDP training set 9 and ITS1 using a custom ITS1 database<sup>4</sup>. *Staphylococcus* and *Malassezia* genera were classified to the species level using pplacer<sup>22</sup> with custom databases.

**Reference-based taxonomic and functional classification.** We compiled a list of complete and draft microbial reference genomes of 2,342 bacterial, 389 fungal, 1,375 viral, and 67 archaeal genomes from the National Center for Biological Information (NCBI, <http://www.ncbi.nlm.nih.gov>), the Human Microbiome Project (HMP, <http://www.hmpdacc.org>), the Saccharomyces Genome Database (SGD, <http://www.yeastgenome.org>), the Fungal Genome Initiative (FGI, <http://www.broadinstitute.org>), FungiDB (<http://fungi.db.org>), and internally sequenced genomes (Supplementary Table 2). Where multiple genomes for a reference were available, we selected complete over draft genomes. Reads not matching hg19 + hg19 rRNA were mapped to this genome collection using bowtie2<sup>23</sup>—very-sensitive parameter retrieving the top 10 hits. Reads mapping to multiple genomes were then reassigned to a 'most likely' genome using Pathoscope v1.0<sup>24</sup>, which uses a Bayesian framework to examine each read's sequence and mapping quality within the context of a global reassignment. Read hit counts were then normalized by genome length and scaled to sum to one. To reduce the likelihood of recovering spurious genomes, we also calculated genome coverage for each genome hit using the genomeCoverageBed tool in the Bedtools suite<sup>25</sup>. For relative abundance and diversity calculations, genomes with coverage  $< 1$  were removed to decrease low-abundance false positives, providing a measure of normalization for sequencing depth.

To assess the accuracy of our taxonomic classifications and our estimation of community diversity, we compared taxonomic assignments of bacteria and fungi to 16S and ITS amplicon results, as well as to the output from a bacterial and archaeal mapping tool, Metaphlan<sup>26</sup>. We observed high correlations extending to the species level for bacterial sequences (Extended Data Fig. 3, Supplementary Tables 2–4). Concordance of non-*Malassezia* fungal species was lower, presumably due to the relative paucity of sequenced fungal genomes. We used the Shannon diversity index as well as species observed for diversity comparisons for bacterial classifications. All taxonomies were reconstructed to the species level, combining hits to multiple strain subtypes. The coverage cutoff of 1 was chosen as an inflection point for species accumulation and as a point of concordance between diversity estimates derived from other approaches.

We characterized the representation of functional gene groups in the skin using the KEGG Orthology gene pathway (KO) and module (MO) annotations<sup>27</sup>, calculating corresponding abundances and coverages using the HMP Unified Metabolic Analysis Network (HUMANN)<sup>28</sup>. We note that functional diversity is probably underestimated in the absence of viral pathways in the KEGG database. We mapped reads to the 2013.10.14 KEGG release using USEARCH v7.0<sup>29</sup> e-value  $< 0.01$ , -accel 0.5 as described<sup>28</sup>. The top 10 hits were then processed with HUMANN v0.99<sup>28</sup>. To define genetic carriage of resistance profiles in the skin, antibiotic resistance genes from the Antibiotic Resistance Genes Database (ARDB)<sup>30</sup> were clustered based on sequence similarity to produce families of unique short sequence markers using ShortBRED (J. Kaminski, N. Segata, E. Franzosa and C. Huttenhower, unpublished). Reads were then mapped to the top marker using USEARCH v7.0, minimum alignment length 20, percent identity 95%. A family (resistance gene) was called present if at least one gene of that family was represented with a non-zero median of all its markers (median number of hits to its markers  $> 0$ ). Each family was normalized by the number of the hits, the marker length, and the length of the original protein sequence. We considered only presence/absence for a more conservative assessment. We note that while antibiotic resistance genes are typically classified with respect to a particular species, from metagenomic data it is difficult to impute an organism of origin because families can be encoded on plasmids (for example, NP\_040465, a tetracycline efflux pump).

**Reference-based strain mapping.** Accurate, *de novo* identification of single nucleotide polymorphisms (SNPs), used in metagenomic strain tracking of high-biomass stool samples, typically requires  $100\times$  coverage for robust identification<sup>31</sup>. Given strain variance due to differential representation and sequencing depth, we developed a reference-based approach, assessing feasibility and accuracy with computational simulations of communities of mixed complexity. For bacteria *Propionibacterium acnes* and *Staphylococcus epidermidis*, we created custom, species-specific reference



databases incorporating all complete and draft genomes present for those species from NCBI, totalling 78 and 61, respectively (Supplementary Table 2). To visualize relationships between the strains, all SNPs identified in core regions were used to create dendrograms with the program PhyML 3.0<sup>32</sup>. Strains were assigned to a subtype based on phylogenetic distance, for example, we defined 12 subtypes for *P. acnes* and 14 for *S. epidermidis*.

For each respective set of reference genomes, we identified first, SNPs unique to each strain in regions shared in all genomes ('core'), and second, larger regions that are partially shared or unique to a strain ('non-core', Supplementary Table 2). We mapped reads to each database using bowtie2 with stringent parameters ( $-\text{score-min L,-0.6,0.006}$ ), allowing zero mismatches and as many hits as genomes in the database. Read assignment using Pathoscope was performed as described, except  $\text{theta\_prior}$ , an option that controls the proportion of non-unique reads that are assigned to a genome, was set to  $10 \times 10^{88}$  (most genomes permitted). Normalization was performed as described above.

Because Pathoscope can reassign reads to closely related genomes rather than an actual target genome that may or may not be present in a sample, we evaluated the ability of Pathoscope to accurately reassign reads to very similar sub-strains by first, assessing sensitivity of complex staggered mixtures of synthetic communities, and second, demonstrating the presence of unique genomic loci that allow discrimination between subtypes. First, synthetic communities were created with 6, 12, or 18 genomes per community, with 50,000, 100,000, or 500,000 reads sampled per genome for an even mix, as well as a staggered community to estimate accuracy in abundance calling. 15 random synthetic communities for each even genome group, and 5 for staggered, were created and mapped to the full genome set. Sensitivity was calculated from the expected versus observed abundances. Second, we identified SNPs unique to each genome in 'core' regions of the genome (defined as shared between all reference genomes in species-specific database) using nucmer<sup>33</sup> and custom scripts. nucmer was also used to identify 'non-core' regions in each of the genomes. Simulated reads were then mapped to strains based upon: (1) consensus SNPs, (2) non-core region variants, or (3) full genomes to identify what variants are shared between sites/individuals. In simulations, core SNPs had the highest sensitivity, but whole genomes, which incorporate both core and non-core elements, were best able to identify closest neighbour strains (Extended Data Fig. 5, Supplementary Table 9). Although we have supported our results using SNPs (Supplementary Table 10), mapping to whole genomes provided clear advantages if an exact reference strain is not present *in vivo*, which is likely given the limited number of fully sequenced genomes. In absence of an exact reference, our approach robustly defines most similar strains based on differentiating genomic features.

**Adaptive iterative de novo assembly.** Assembly efficacy varies depending on the site's unique features of community complexity, typically defined by microenvironment, and sequencing depth, which is affected by biomass and human DNA admixture. To optimize assembly parameters, individual samples were assembled using a wide *k*-mer range in Velvet<sup>34</sup>, and contigs greater than 300 bp in length were analysed. To examine assembly efficacy, reads were remapped to assemblies using bowtie2—sensitive. 'Adaptive' denotes that each sample was assembled using *k*-mers ranging from 37–69. A quality score was calculated using % paired or singleton reads realigning to the assembly, the number of bases incorporated into the assembly, and number of contigs >300 bp. The assembly with the highest quality score was used for subsequent analysis. 'Iterative' denotes subsequent steps in which unaligned reads from remapping were then pooled to improve recovery of rare genes that may represent genomes unique to an individual. We found that pooling by individual produced higher quality assemblies than pooling by site (Supplementary Table 17). This observation supported our insight that while site can shape the major features of a community, species and strains are shared within an individual. To improve assembly quality and reduce computational burden, digital normalization<sup>20</sup>, which reduces error by removing redundant data and performs similarly to non-normalized data (Extended Data Figure 10c), was applied on pooled samples before assembly. We used two-pass normalization to  $20\times$  then  $5\times$  with variable coverage and assembled with adaptive *k*-mer selection. Finally, unaligned reads from pooled individual assemblies were pooled and subsampled 1:10 before normalization and variable assembly.

To create a multi-kingdom skin microbial gene catalogue, genes were predicted from contigs using two models, MetaGeneMark<sup>35</sup>, which incorporates multiple bacterial models, and Augustus<sup>36</sup> with a *Ustilago maydis* model as a phylogenetically near neighbour to *Malassezia*, the most predominant skin fungi. To account for cases where both fungal and bacterial genes were called for the same contig, we adopted a filtering methodology by which each contig was assigned to a kingdom using blastn against our microbial database, or where no blastn hit was available, a blastx against nr using USEARCH. Discordant calls not resolved by blastn/x filtration were marked ambiguous or assigned to whichever caller generated a prediction. A non-redundant catalogue was constructed using UCLUST with sequence

identity cut-off of 0.95 and a minimum coverage cutoff of 0.9 for shorter sequences. This final catalogue contained 5,922,920 putative bacterial and fungal genes.

During this process, we also observed that many short contigs (<1,000 bp) produced no putative genes. To circumvent losing partial genes or genes unidentifiable by our prediction models, we revised our gene catalogue to first retrieve contigs <1,000 bp, then call genes on contigs >1,000 bp as previously described. To assess the abundance of genes, reads were aligned to the gene catalogue with Bowtie2—sensitive and counts per gene were normalized by length.

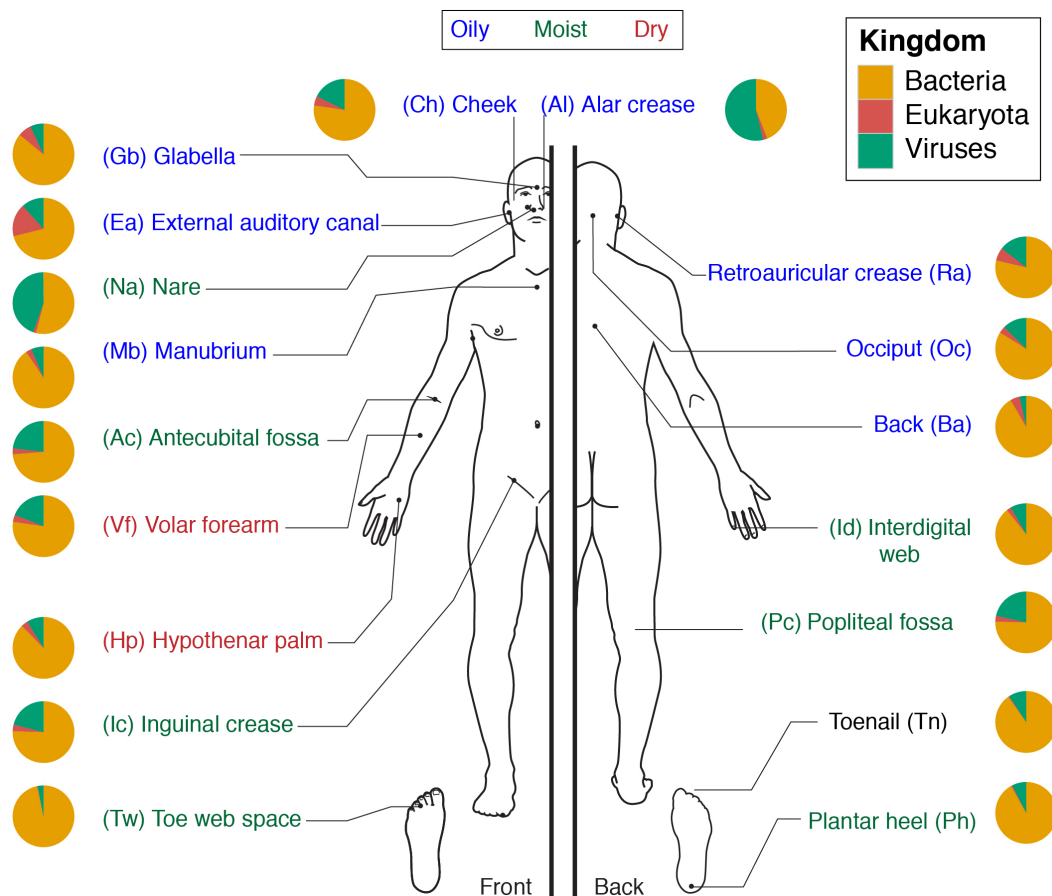
Putative metagenomic clusters, based on covariance of gene abundances across samples, were formed as described<sup>10</sup>. Genes from the same genome are assumed to co-vary in relative abundance across subjects due to physical linkage; therefore such clusters can serve as a proxy for unknown organisms or known organisms with variable gene content. We clustered gene abundances across samples, grouped by site characteristic both to improve segregation of clusters and reduce computational burden. To reduce false positives and computational complexity, we required genes to be present in at least 20% of samples for a given site characteristic. The abundances of these genes across samples were then clustered using the Markov clustering algorithm implemented in MCL<sup>37</sup> with a Spearman correlation coefficient of 0.85 and inflation parameter set to 2. Cluster parameters varying presence to 40% presence across samples, correlation coefficients to 0.80 and 0.90, and inflation parameters of 4 produced similar results. For toenail, 40% presence and clustering at 80% was performed due to computational limitations imposed by site complexity. Clusters were taxonomically annotated by blastx-ing each gene in a cluster to nr as previously described, and as a strict requirement against false binning, clusters with at least 50% of genes mapping to the same phylogenetic group at the species, genus, and/or family level were retained as a metagenomic 'cluster'. Clusters with the same consensus taxonomy were merged at the genus and species level; family level analysis showed minimal improvements in consensus (Supplementary Table 18). Because a typical microbial genome contains thousands of genes, we speculate that many of these represent gene fragments that did not pass our stringent redundancy thresholds. While our variable sequencing depth likely precludes recovery of complete genomes from such a metagenomic linkage analysis, we identified large clusters of taxonomically related groups of covarying genes for both characterized and uncharacterized species.

**Statistical analysis.** All statistical analyses were performed in the R software. Data are represented as mean  $\pm$  standard error of the mean unless otherwise indicated. For all boxplots, black centre lines represent the median and box edges the first and third quartiles. 'e' in scientific notation refers to  $10 \times$ , for example,  $10e5$  represents  $10 \times 10^5$ . Spearman correlations ( $\rho$ ) of non-zero values were used for all correlation coefficients. The nonparametric tests Wilcoxon rank-sum and Kruskal–Wallis were used to determine statistically significant differences between microbial populations, and to identify significant inter-category comparisons, we used a post-hoc multiple comparison test, implemented by the *kruskalmc* test in the *pgirmess* package. Unless otherwise indicated, *P* values were adjusted for multiple comparisons using the *p.adjust* function in R using method = "fdr"<sup>38</sup>. Statistical significance was ascribed to an alpha level of the adjusted *P*-values  $\leq 0.05$ . Site characteristics were treated as separate groups where indicated based on spatial physiological differences between these different body niches<sup>2</sup>. Similarity between samples was assessed using the Yue–Clayton theta similarity index<sup>39</sup> with relative abundances of species, sub-strains, or shared genomic variants. The theta coefficient assesses the similarity between two samples based on (1) number of features in common between two samples, and (2) their relative abundances with  $\theta = 0$  indicating totally dissimilar communities and  $\theta = 1$  identical communities. To avoid repeated measures, samples belonging to an individual were averaged before statistical comparisons between site characteristic when using summary metrics such as means, diversity, or theta indices.

Supervised random forest models to identify discriminatory taxa and modules was implemented with the *randomForest* package in R<sup>40</sup>. This analysis was enabled by our multi-site sampling strategy, as using a single or few sites lacks statistical power to detect low abundance features. Mean decrease in accuracy denotes the normalized difference in the classification accuracy when that variable is included versus when data are randomly permuted, that is, to what degree inclusion of this predictor in the model reduces classification error. Model accuracy was calculated using the out-of-bag (oob) error estimate, which is an approximation of how frequently an individual is misclassified.

19. Jumpstart Consortium Human Microbiome Project Data Generation Working Group. Evaluation of 16S rDNA-based community profiling for human microbiome research. *PLoS ONE* **7**, e39315 (2012).
20. Howe, A. C. *et al.* Tackling soil diversity with the assembly of large, complex metagenomes. *Proc. Natl Acad. Sci. USA* **111**, 4904–4909 (2014).
21. Schloss, P. D. *et al.* Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**, 7537–7541 (2009).

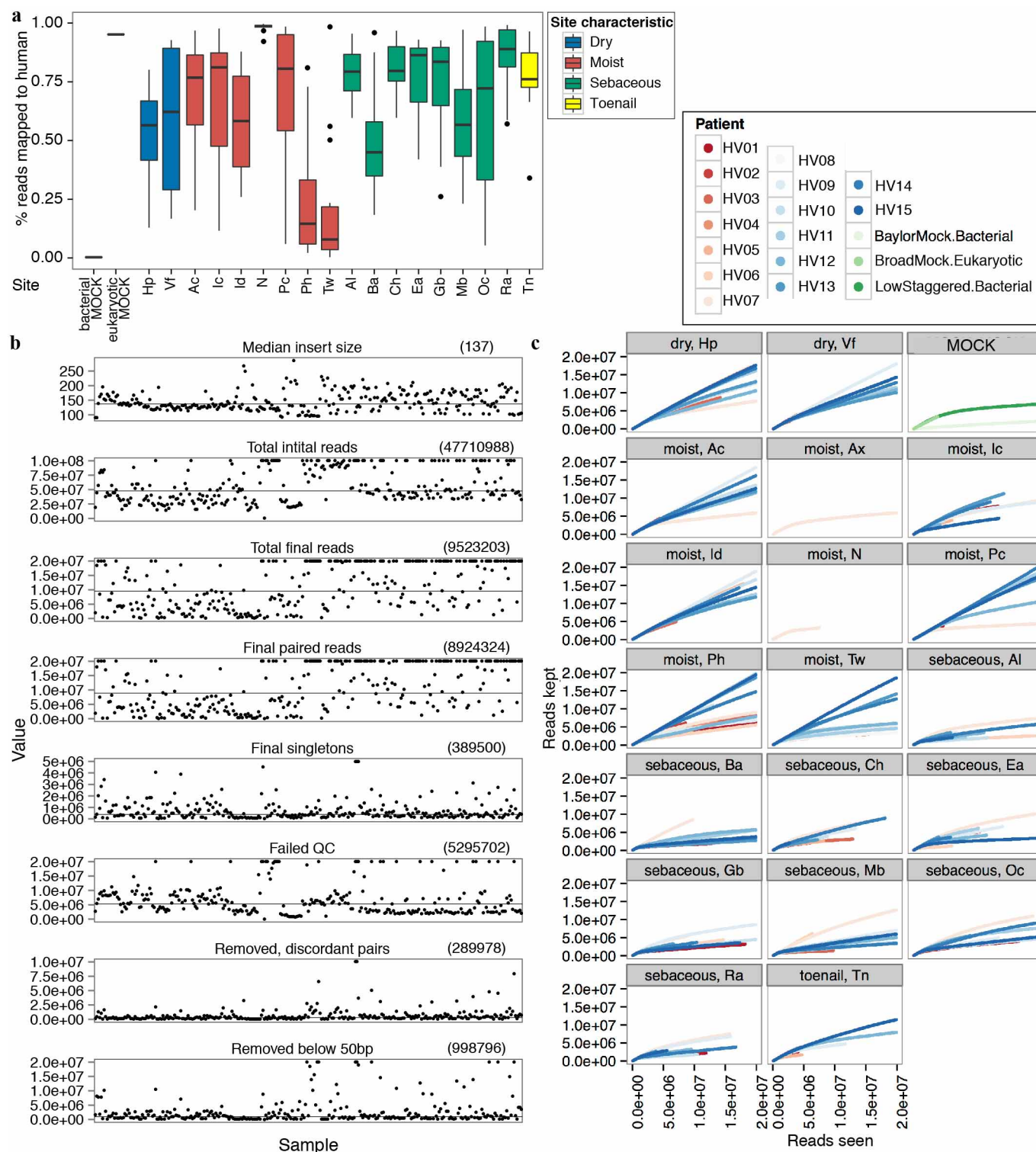
22. Matsen, F. A., Kodner, R. B. & Armbrust, E. V. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* **11**, 538 (2010).
23. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
24. Francis, O. E. *et al.* Pathoscope: Species identification and strain attribution with unassembled sequencing data. *Genome Res.* **23**, 1721–1729 (2013).
25. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
26. Segata, N. *et al.* Metagenomic microbial community profiling using unique clade-specific marker genes. *Nature Methods* **9**, 811–814 (2012).
27. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
28. Abubucker, S. *et al.* Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput. Biol.* **8**, e1002358 (2012).
29. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
30. Liu, B. & Pop, M. ARDB—Antibiotic Resistance Genes Database. *Nucleic Acids Res.* **37**, D443–D447 (2009).
31. Schloissnig, S. *et al.* Genomic variation landscape of the human gut microbiome. *Nature* **493**, 45–50 (2013).
32. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
33. Delcher, A. L., Phillippy, A., Carlton, J. & Salzberg, S. L. Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Res.* **30**, 2478–2483 (2002).
34. Namiki, T., Hachiya, T., Tanaka, H. & Sakakibara, Y. MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. *Nucleic Acids Res.* **40**, e155 (2012).
35. Zhu, W., Lomsadze, A. & Borodovsky, M. *Ab initio* gene identification in metagenomic sequences. *Nucleic Acids Res.* **38**, e132 (2010).
36. Stanke, M., Schöffmann, O., Morgenstern, B. & Waack, S. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics* **7**, 62 (2006).
37. van Dongen, S. & Abreu-Goodger, C. in *Bacterial Molecular Networks* (eds Helden, J., Toussaint, A. & Thieffry, D.) *Methods in Molecular Biology* Vol. 804, pp. 281–295 [http://link.springer.com/protocol/10.1007/978-1-61779-361-5\\_15](http://link.springer.com/protocol/10.1007/978-1-61779-361-5_15) (Springer, 2012).
38. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
39. Yue, J. C. & Clayton, M. K. A similarity measure based on species proportions. *Comm. Stat. Theory Methods* **34**, 2123–2131 (2005).
40. Liaw, A. & Wiener, M. Classification and regression by randomForest. *R News* **2**, 18–22 (2002).



**Extended Data Figure 1 | The 18 selected skin sites and their location on the human body.** These sites represent three microenvironments: sebaceous (blue), dry (red), and moist (green). Toenail (black) is a site that does not fall

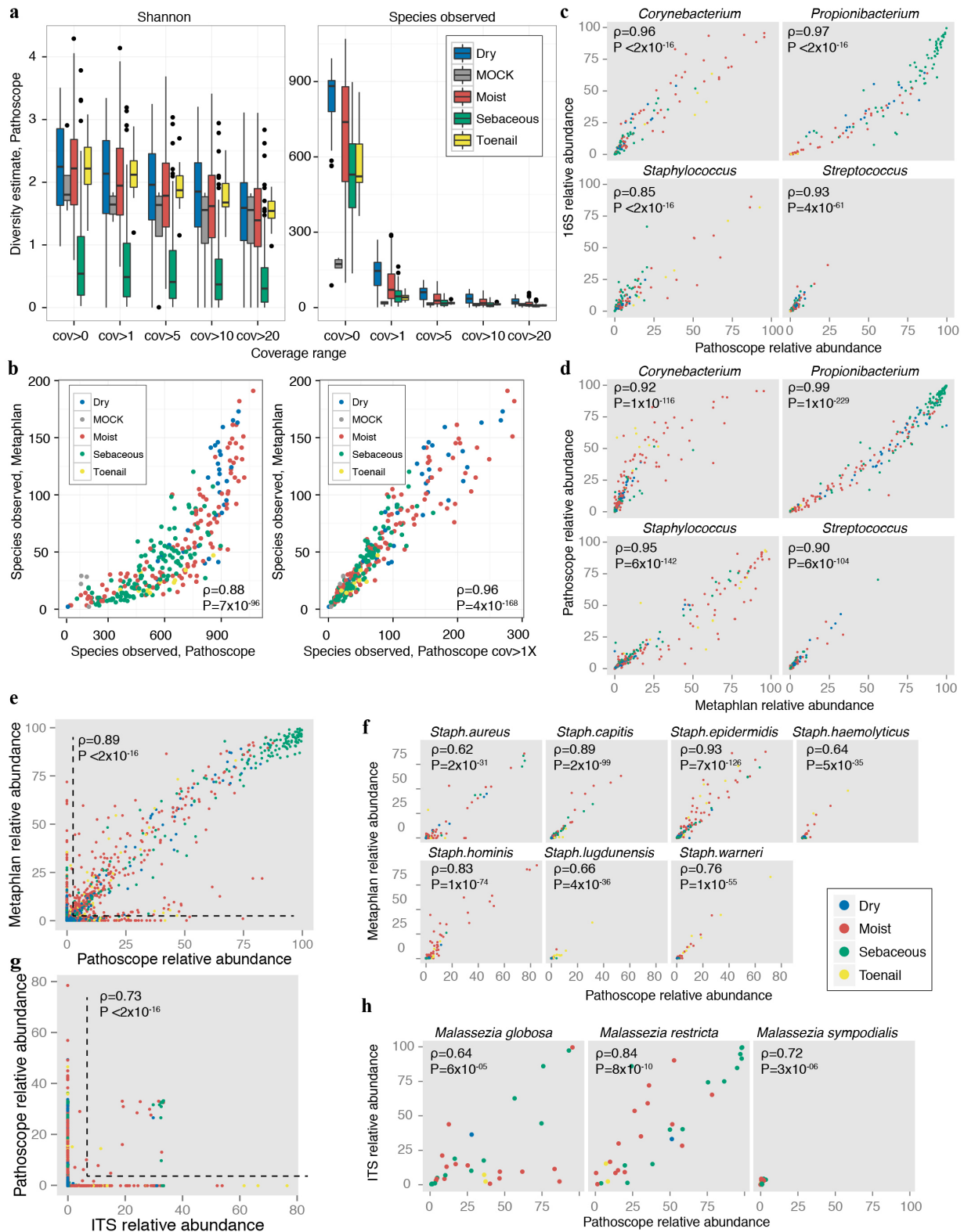
under these major microenvironments and is treated separately. Pie charts represent consensus relative abundance of the kingdoms Bacteria, Eukaryota (Fungi), and virus from multi-kingdom mapping.





**Extended Data Figure 2 | Per-sample read statistics.** Additional samples (bacterial and eukaryotic mock communities) are shown. **a**, Boxplots (line indicates median; boxes represent first and third quartiles) show, for each site, % reads mapping to human hg19 that are discarded before analysis. Sites are coloured by site characteristic. **b**, Samples are ordered by label. Lines indicate

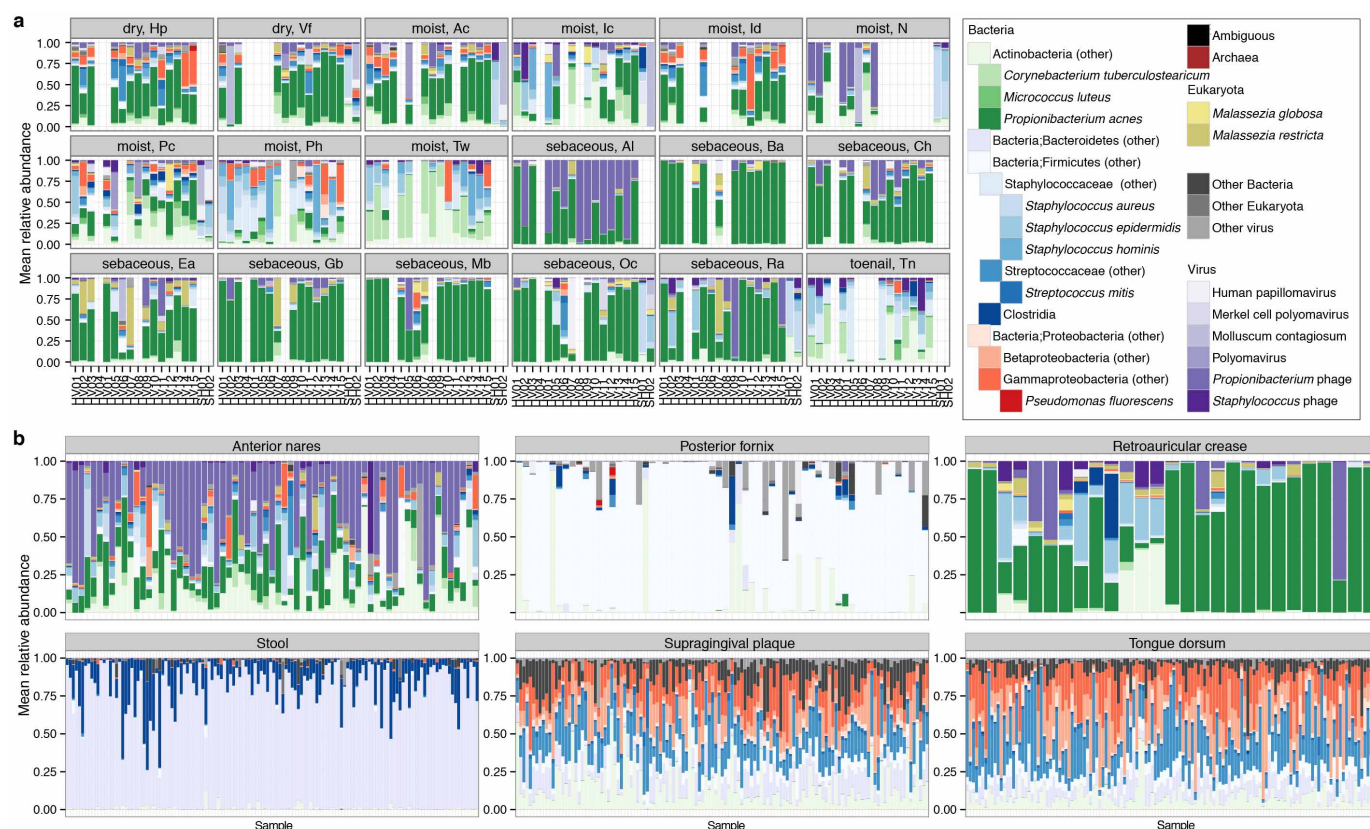
the median value for that statistic; value is in parenthesis. **c**, Estimate of sequencing coverage. Reads seen is the number of reads in a sample sampled. Reads are then split into 20-mers, compared to a  $k$ -mer coverage table and kept only if the median  $k$ -mer coverage is below 20 $\times$ . Curves are grouped by site, coloured by individual as indicated.



### Extended Data Figure 3 | Validation of taxonomic classifications.

**a**, Bacterial sample community diversity as a function of genome coverage for two diversity metrics, the Shannon index that measures the richness and evenness of the community (left), and number of species observed (right). Genome coverage is defined as for each genome hit, the % of genome covered by reads. Boxplots show the range of diversity values for all samples, segregated by microenvironment. Black lines indicate median; boxes represent first and third quartiles. As coverage cut-offs increase, diversity estimates drop sharply. **b**, Comparisons of bacterial community diversity for Metaphlan-derived classifications versus custom bacterial Pathoscope-derived classifications. Each

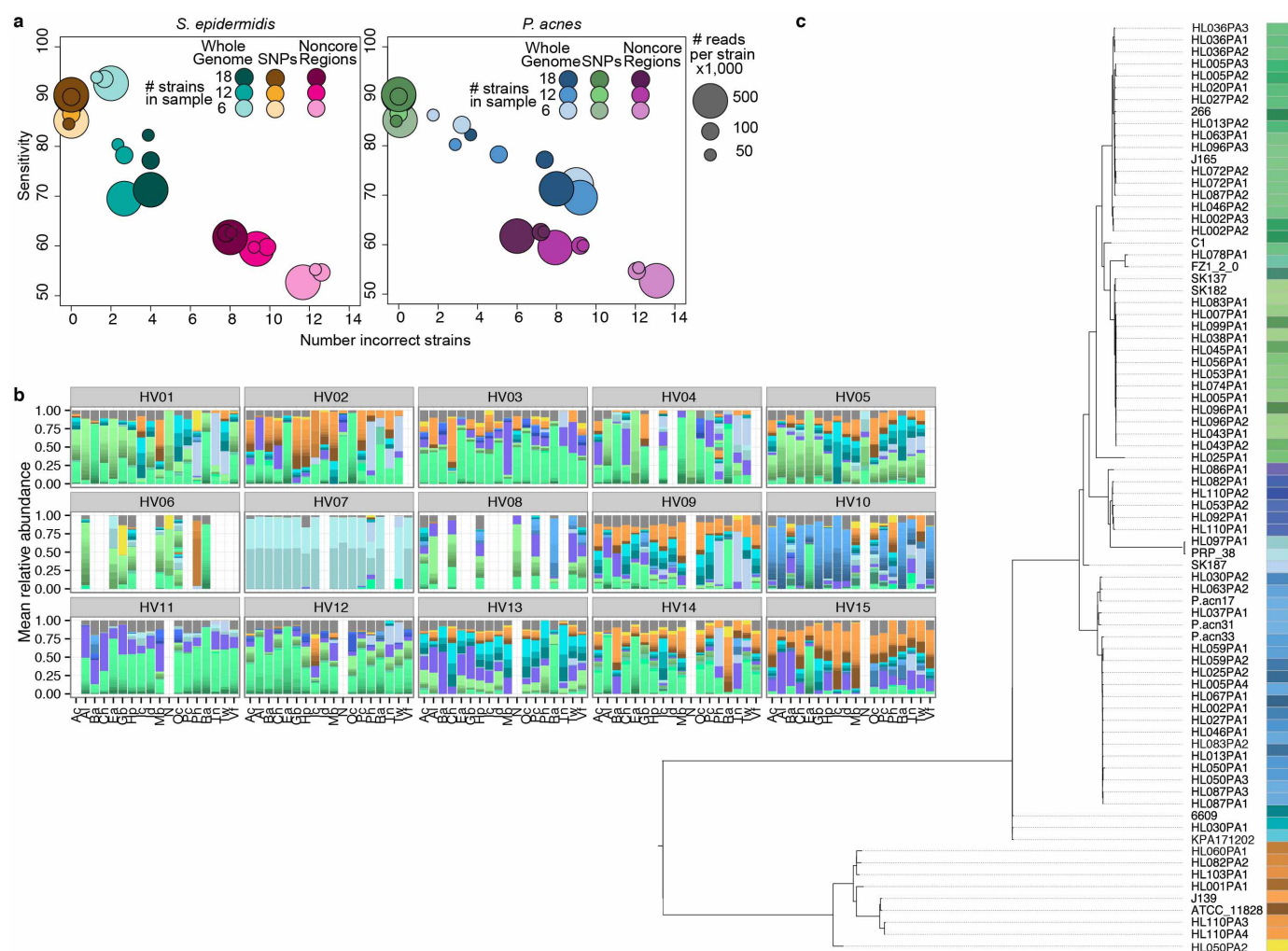
point represents a different sample, coloured by microenvironment. With no coverage cut-offs (left), Pathoscope may overestimate diversity, which is reduced by setting a minimum 1× coverage requirement. Spearman correlation ( $\rho$ ) and corresponding  $P$  values are shown. Pathoscope-derived relative abundances versus relative abundances derived from **c**, 16S amplicon sequencing, **d**, Metaphlan genus-level, **e**, Metaphlan-species level ( $\rho$  and  $P$  value are calculated for non-zero abundance taxa), **f**, Metaphlan, staphylococcal species, **g**, ITS1 amplicon sequencing, genus ( $\rho$  and  $P$  value are calculated for non-zero abundance taxa), and **h**, ITS1 amplicon sequencing, *Malassezia* species.



**Extended Data Figure 4 | Full taxonomic classifications for all healthy volunteers (HV), all sites.** To aid visualization of site- and individual-specific similarities, samples are grouped by site/microenvironment for each individual. Relative abundances of the most abundant skin taxa for each super-kingdom are shown. **b**, Taxonomic re-classification of major sites sampled by the

Human Microbiome Project. Samples are from the anterior nares and retroauricular crease (skin), tongue dorsum and supragingival plaque (oral), stool, and posterior fornix (vaginal). Relative abundances of the most abundant taxa for each kingdom in the skin, for comparison, are shown.

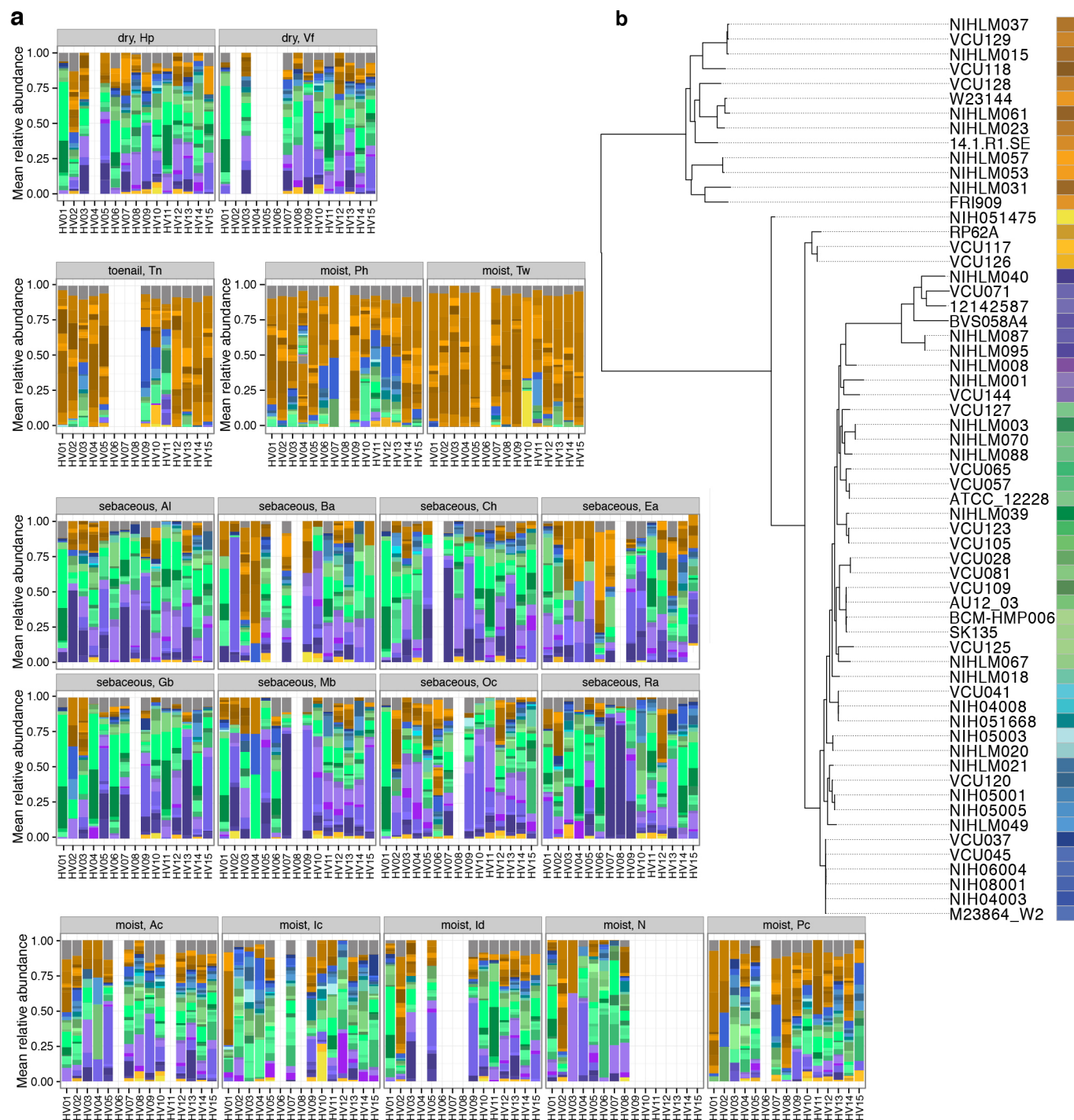




**Extended Data Figure 5 | Strain-level classification based on reference genomes show sub-species heterogeneity for dominant skin taxa.**

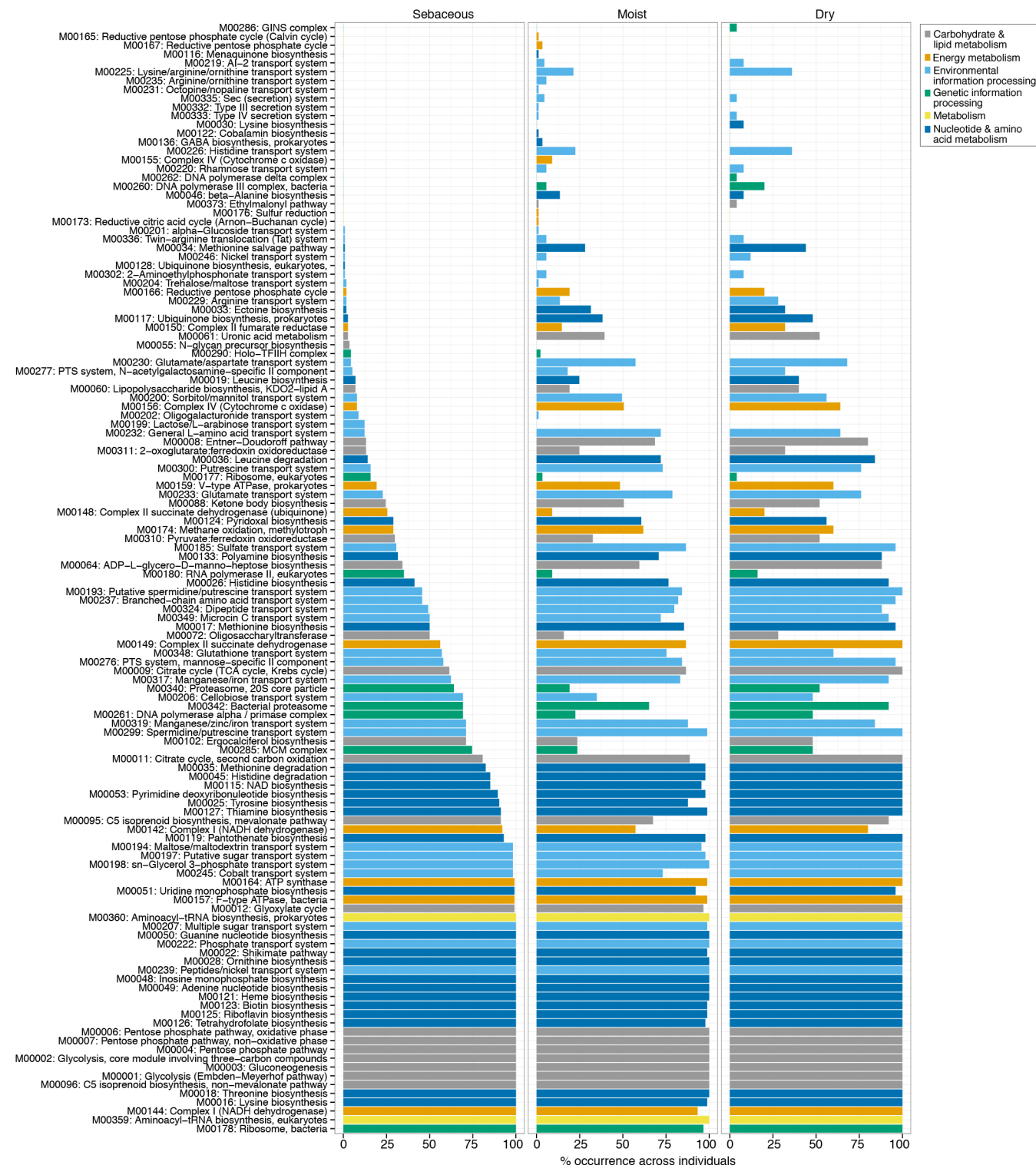
**a**, Simulations to assess sensitivity of Pathoscope-based mapping to SNPs, non-core regions, or whole genomes. Synthetic communities were created with 6, 12, or 18 genomes per community. Sizes of circles reflect the number of reads sampled from each genome, for example, 50,000, 100,000, or 500,000 reads per genome. 15 random synthetic communities for each genome group were created and mapped to SNPs, non-core regions, or the full genome set. Sensitivity is calculated from the expected versus the observed abundances.

**b**, Full strain-level assignments for samples with relative abundances of closest related *Propionibacterium acnes* strains, by individual. **c**, Dendrograms of strain similarity. Trees were generated using core SNPs; genomes were aligned with nucmer to identify core regions, and then SNPs within these core regions were identified by calculating all pairwise differences between genomes. Bar of colours indicates delineations of subtypes where phylogenetically more similar genomes are in similar colours; for example, we defined 12 subtypes for *P. acnes*.



**Extended Data Figure 6 | Strain-level classification for *Staphylococcus epidermidis*.** **a**, Full strain-level assignments for samples by

microenvironment. **b**, Description is as in Extended Data Fig. 5c. We defined 14 subtypes for *S. epidermidis*.



**Extended Data Figure 7 | Full version of coreness of different module categories across skin microenvironment.** A module is defined as core if occurring in  $>2/3$  of samples for that class. Major KEGG module descriptors

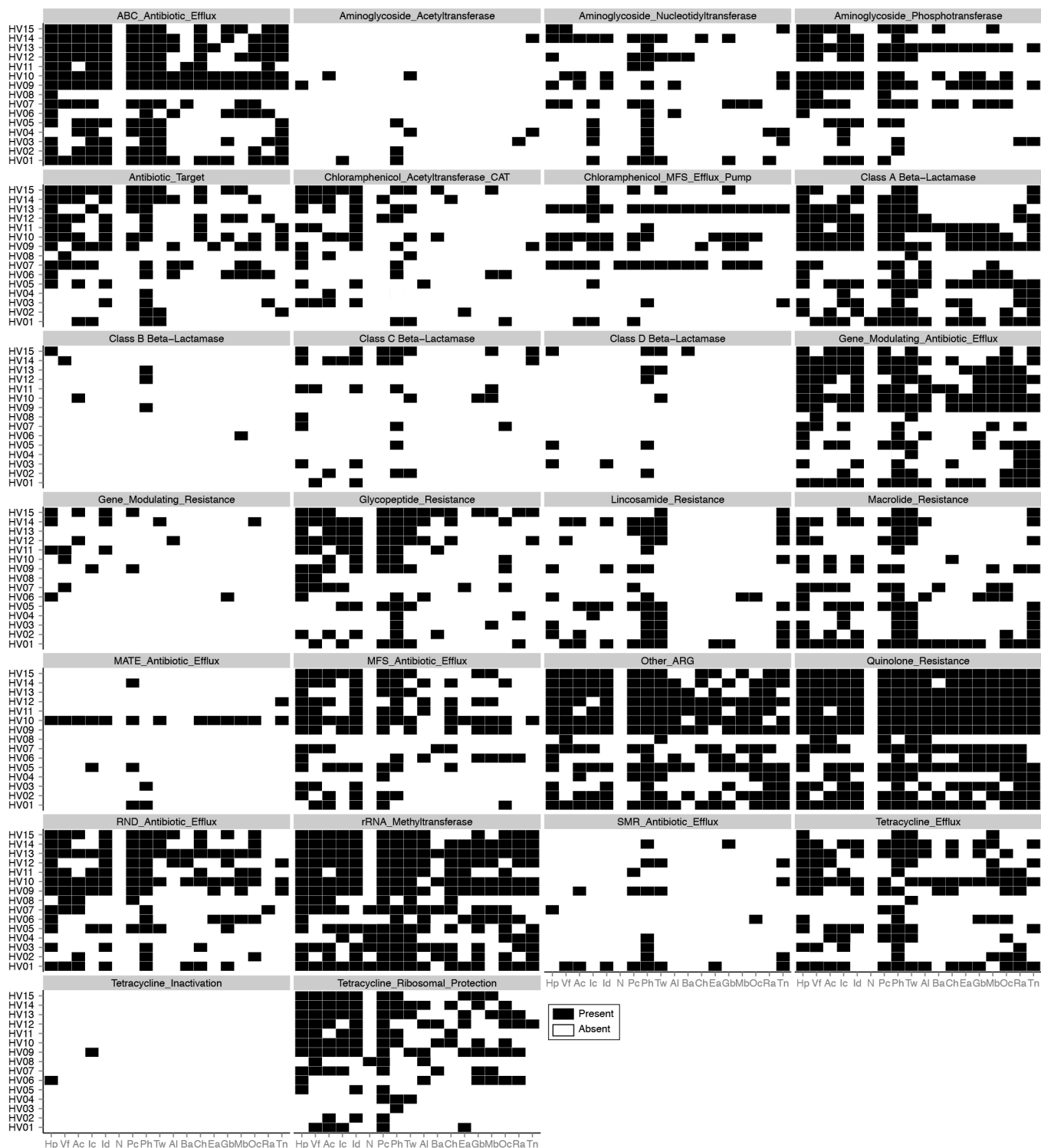
are shown in the different colours. Height of bars reflects the proportion of samples that a module occurs in.





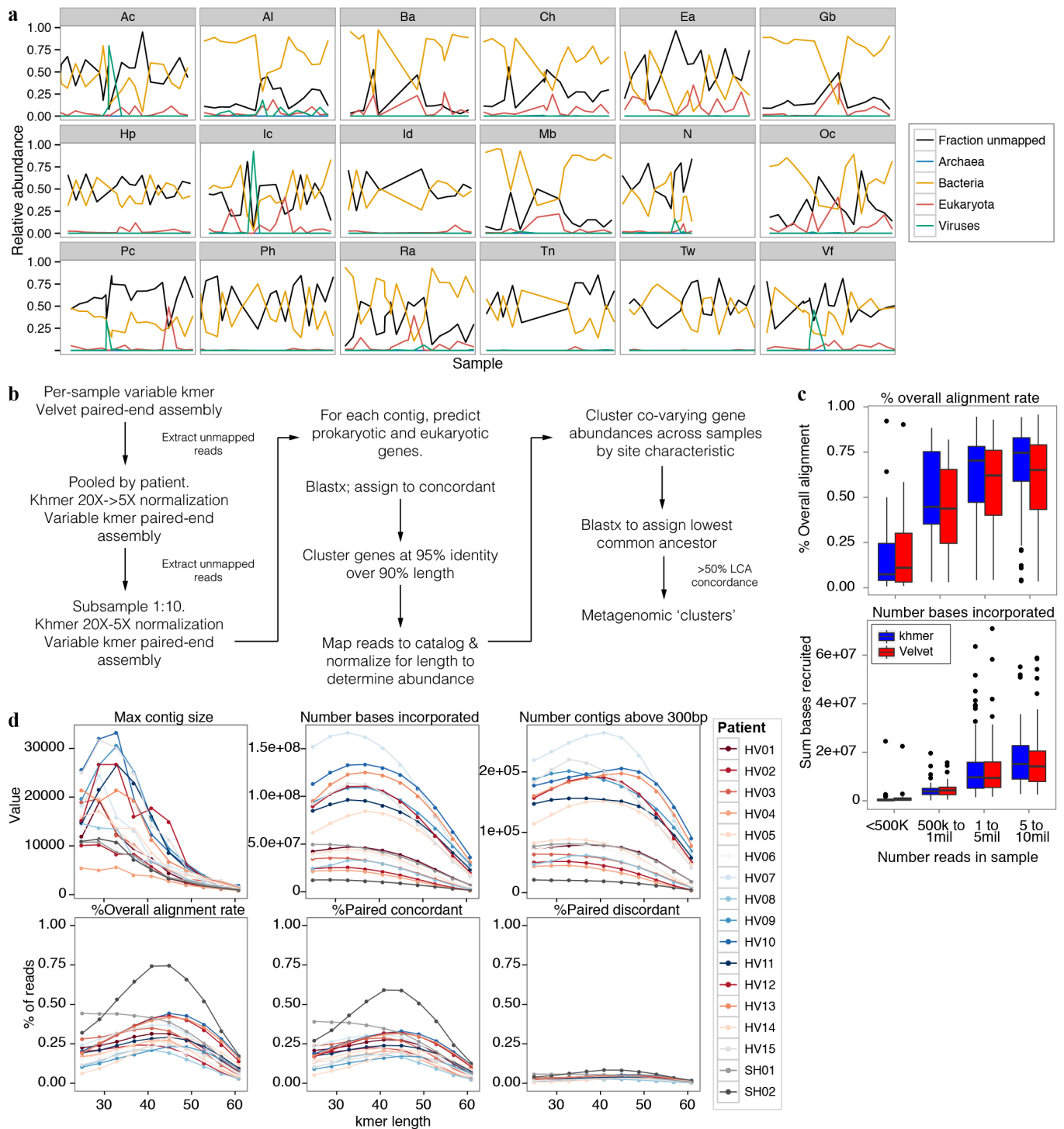
**Extended Data Figure 8 | Correlation analysis of module abundance with species abundance to infer a module's taxonomic origin.** Spearman correlation ( $\rho$ ) was calculated with corresponding  $P$  value for taxa with relative abundance  $>0.5\%$  and modules with greater than  $0.05\%$  relative abundance.

*Coryn.*, *Corynebacterium*. **a**, Unsupervised clustering of correlation coefficients. Species from the same genera clustering together may suggest a shared contribution of a pathway. **b**, Most significantly correlated taxa; colours represent broad KEGG classes. Adjusted  $P < 2 \times 10^{-16}$ .



**Extended Data Figure 9 | Antibiotic resistance profiles in the skin.** Reads were mapped to a short marker database consensus created from the ARDB database, which catalogues publicly available resistance genes. Genes are

grouped into broad resistance classes; a resistance category is called present (black; absent = white) if at least one gene from its family is present.



### Extended Data Figure 10 | Reference-free analysis of skin metagenome with adaptive iterative assembly, gene catalogue, and metagenomic clusters.

**a**, Tracking unclassified reads. Fraction unmapped reads refers to the fraction of total reads passing quality control that do not map to the major super kingdoms Archaea, Bacteria, Eukaryota, and viruses. Samples are ordered by label and are divided by site. **b**, Assembly, gene-calling, and clustering workflow. **c**, Assembly efficacy varies significantly by *k*-mer depending on the site's unique features of community complexity and sequencing depth, which is most affected by that site's human DNA admixture. Assembly statistics are shown for samples pooled by individual, which produced higher quality assemblies than pooling by site. Because of large pool size, khmer digital

normalization was used before Velvet assembly. % overall alignment rate indicates the total % of reads that map back to that sample's assembly for each *k*-mer. % paired concordant indicates the fraction paired reads (of overall, not of % paired) in which both pairs of a mate map back to an assembly; discordant is where one mate of a pair does not map, or maps to a different contig. Contigs are then assessed by the maximum assembly size, the number of bases that are used in the assembly, and the number of contigs above a threshold of 300 bp. **d**, Effect of khmer digital normalization on individual sample assembly. Digital normalization + Velvet assembly performs similarly to Velvet assembly alone.



# HCN ice in Titan's high-altitude southern polar cloud

Remco J. de Kok<sup>1,2</sup>, Nicholas A. Teanby<sup>3</sup>, Luca Maltagliati<sup>4</sup>, Patrick G. J. Irwin<sup>5</sup> & Sandrine Vinatier<sup>4</sup>

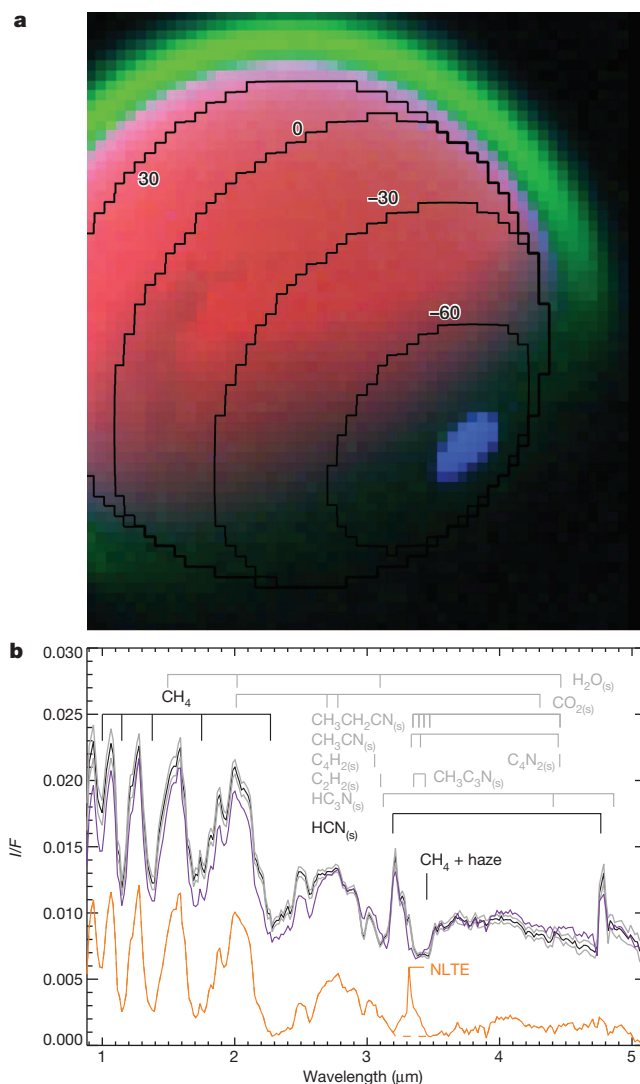
Titan's middle atmosphere is currently experiencing a rapid change of season after northern spring arrived in 2009 (refs 1, 2). A large cloud was observed<sup>3</sup> for the first time above Titan's southern pole in May 2012, at an altitude of 300 kilometres. A temperature maximum was previously observed there, and condensation was not expected for any of Titan's atmospheric gases. Here we report that this cloud is composed of micrometre-sized particles of frozen hydrogen cyanide (HCN ice). The presence of HCN particles at this altitude, together with temperature determinations from mid-infrared observations, indicate a dramatic cooling of Titan's atmosphere inside the winter polar vortex in early 2012. Such cooling is in contrast to previously measured high-altitude warming in the polar vortex<sup>1</sup>, and temperatures are a hundred degrees colder than predicted by circulation models<sup>4</sup>. These results show that post-equinox cooling at the winter pole of Titan is much more efficient than previously thought.

In May 2012, a large cloud-like structure was identified above Titan's dark southern pole by Cassini's Imaging Science Subsystem (ISS)<sup>3</sup>. Ever since, it has been seen at very high altitudes ( $\sim 300$  km) and high southern latitudes, at all visible and near-infrared wavelengths. Clouds require temperatures cold enough for atmospheric gases to reach saturation. Hence, clouds on Titan have previously been found near the tropopause and lower stratosphere, where the atmosphere is coldest<sup>5–8</sup>. Instead of a temperature minimum, a temperature maximum was present before 2012 at the altitudes and latitudes where the high-altitude ISS cloud is seen<sup>1,9</sup>. Such high temperatures precluded the condensation of any of Titan's known trace gases. The presence of a cloud at this location is therefore highly unexpected.

We analysed near-infrared spectra of the high-altitude cloud from Cassini's Visual and Infrared Mapping Spectrometer (VIMS) to constrain its composition and optical thickness. Near-infrared wavelengths are sensitive to vibrational bands of solids and liquids and can therefore be used to identify the cloud composition. We have averaged the spectra of the high-altitude cloud in 13 similar VIMS image cubes from 29 November 2012, with pixel scales between 89 and 135 km (Fig. 1a), to obtain a high signal-to-noise near-infrared reflectance spectrum of the cloud. This spectrum indeed shows two large spectral features (Fig. 1b), which are not present in cloudless regions. The spectral features coincide exactly with the features expected from HCN ice<sup>10,11</sup> and are detected at a level at least 15 times the standard deviation of reflectance of a single pixel in a single image. Other possible condensates clearly do not match the spectral features seen in the data. We fitted the reflectance spectrum of the cloud and found excellent agreement with a simple model that includes scattering by an optically thin cloud that is composed of HCN ice particles with a radius between 0.6 and 1.2  $\mu\text{m}$  (Fig. 1b). This is the first strong evidence for HCN condensation in Titan's stratosphere—we note an earlier tentative identification of HCN ice<sup>12</sup>, and further indirect evidence<sup>13</sup>.

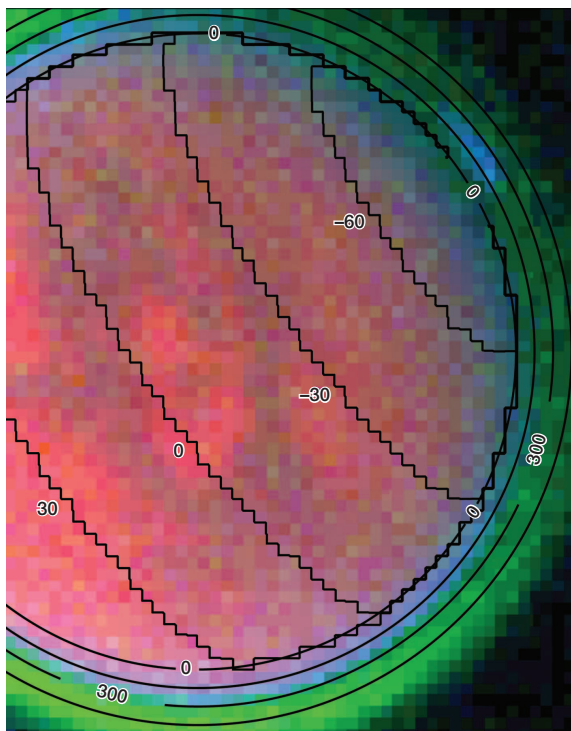
We obtained an estimate of the optical thickness of the cloud from a VIMS image cube from 7 June 2012, which has a pixel scale of 91 km, where the cloud was seen at the limb of Titan (Fig. 2). We determined the cloud top to be located at an altitude of  $300 \pm 70$  km, based on the fact that the highest HCN cloud pixels intersect the line of 300 km altitude

through their centres. This estimate is consistent with results from the ISS instrument<sup>3</sup>, but has a greater uncertainty due to the larger pixel size of VIMS. In Fig. 2, sunlight travels along a slanted path through the



**Figure 1 | Identification of HCN ice in VIMS observations.** **a**, A single false-colour VIMS image from 29 November 2012 indicating the illuminated surface (wavelength 1.07  $\mu\text{m}$ , shown red), non-LTE emission (3.33  $\mu\text{m}$ , green), and an HCN ice feature (3.21  $\mu\text{m}$ , blue). Contours show surface latitudes. Solar illumination is from the upper left. **b**, Mean spectrum away from the cloud (orange, dashed line indicates how the non-LTE (NLTE) emission is removed for the fitting procedure) and within the cloud (black, offset by 0.005 for clarity; grey lines indicate  $\pm 1$  s.d. from a single pixel). Wavelengths of HCN ice features, and of features from other possible condensates<sup>10,11,25,26</sup>, are indicated. A fit to the cloud spectrum is plotted in purple.

<sup>1</sup>Leiden Observatory, Leiden University, Postbus 9513, 2300 RA, Leiden, The Netherlands. <sup>2</sup>SRON Netherlands Institute for Space Research, Sorbonnelaan 2, 3584 CA Utrecht, The Netherlands. <sup>3</sup>School of Earth Sciences, University of Bristol, Wills Memorial Building, Queen's Road, Bristol BS8 1RJ, UK. <sup>4</sup>LESIA-Observatoire de Paris, CNRS, UPMC Université Paris 06, Université Paris-Diderot, 5 place Jules Janssen, F-92195 Meudon, France. <sup>5</sup>Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Clarendon Laboratory, Parks Road, Oxford OX1 3PU, UK.



**Figure 2 | Cloud observed at Titan's limb.** False-colour VIMS image from 7 June 2012 showing Titan's polar cloud at the limb, with altitudes (km) and surface latitudes (degrees) indicated. Colours are as in Fig. 1a. Illumination is from behind the observer. The cloud is seen at top right. The blue/purple colour of the entire limb is caused by Titan's visible disk being larger at  $3.2\ \mu\text{m}$  than at  $1.07\ \mu\text{m}$ . In this image, the cloud reflects less light compared to the rest of the disk than in Fig. 1a, making the limb relatively more blue.

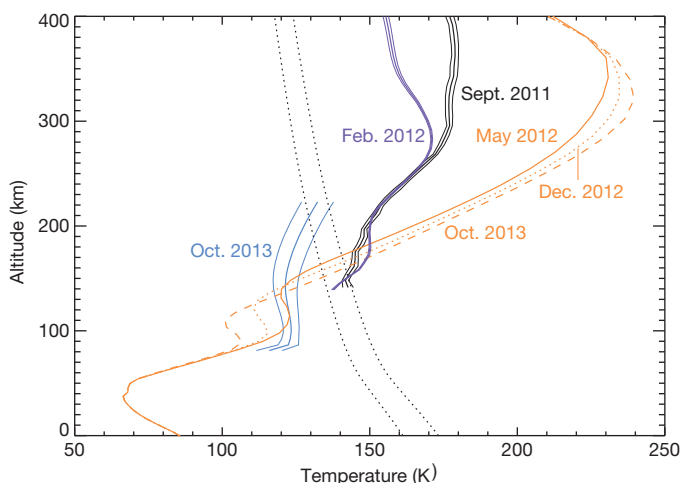
atmosphere, before being scattered back by the cloud to the Cassini spacecraft along almost the same path. In this geometry, the reflectance spectrum is dominated by the reflecting properties of the cloud and the transmission of the atmosphere along the slanted path. Unlike the geometry in Fig. 1a, there is little background contribution from lower altitudes, making it easier to assess the optical thickness of the cloud. At a wavelength of  $2.7\ \mu\text{m}$ , the atmosphere is expected to be practically transparent at an altitude of 300 km, even for slanted paths<sup>14,15</sup>, and almost all signal will be caused by scattering from the cloud. Using the single-scattering approximation for low optical thicknesses, the reflectance can be assumed to be the product of the optical thickness of the slanted path, the single-scattering albedo of the particles, and the phase function at the scattering angle, divided by four. The measured reflectance at  $2.7\ \mu\text{m}$  of  $0.0028 \pm 0.0002$  then directly relates to a slanted-path optical thickness of  $\sim 0.09 \pm 0.006$ , assuming micrometre-sized HCN particles. Since the single-scattering albedo and phase function at  $2.7\ \mu\text{m}$  do not change rapidly with particle size, the slanted-path optical thickness is accurate within a factor of two for the particle size range 0.6– $1.2\ \mu\text{m}$ . If the vertical extent of the cloud is ten times smaller than the length of the slanted path, this translates to a vertical optical thickness of between 0.01 and 0.07 for particle sizes between 0.6 and  $1.2\ \mu\text{m}$  at a wavelength of  $0.9\ \mu\text{m}$ ; results at this wavelength can be compared with the analysis of ISS data.

Although the slanted-path optical thickness can be measured relatively well, an estimate of the particle density requires knowledge of the path length through the cloud, and the exact pressure of the cloud. We perform an order-of-magnitude estimate here using conservative errors on the pressure and path length. Assuming a slanted-path length of tens to hundreds of kilometres (depending on the three-dimensional extent of the cloud) and a pressure of 0.1–0.5 mbar (corresponding to altitudes between 200 and 300 km), this slant optical thickness translates to a particle density of the order of  $10^4$ – $10^5$  particles per gram of gas for

micrometre-sized particles. This particle density falls well within the expected particle density range of micrometre-sized HCN particles with a reasonable downward wind speed of the order of  $0.1$ – $1\ \text{mm s}^{-1}$  (ref. 16). This wind speed is of the same order as the previously inferred downward velocities at the south pole during southern autumn<sup>1,2</sup>. Hence, the optical thickness of the cloud is within the range of that expected for HCN particles, if the temperature is cold enough for them to exist.

At the same time as the appearance of the cloud in ISS and VIMS observations, a condensate feature also appeared at the south pole in far-infrared spectra from Cassini's Composite InfraRed Spectrometer (CIRS)<sup>17</sup>. This feature had been previously observed at the northern pole by Voyager<sup>18</sup> and by Cassini since its arrival in 2004<sup>12,19,20</sup>. It could potentially be linked to the condensation of HCN, since it appeared at a location where HCN condensation was expected to give a large cloud signature<sup>16</sup>. Unfortunately, the available CIRS observations to date cannot constrain the altitude of the far-infrared condensate feature in the south, making it impossible to firmly establish a connection between this feature and the HCN cloud discussed in the present paper. A limb scan with high spatial resolution, which would resolve this issue, is not planned until at least 2015 due to the orbital geometry of Cassini.

Although the VIMS observations of the high-altitude cloud are entirely consistent with the presence of HCN ice particles, the required low temperatures of  $\sim 125\ \text{K}$  are unexpected. The best way to study the temperature of Titan's middle atmosphere is to use mid-infrared spectra from CIRS. Assuming methane is uniformly mixed in the stratosphere, temperatures can be derived from emission of the  $\nu_4$  methane band between 7 and  $8\ \mu\text{m}$ . CIRS measurements of Titan's limb can derive a spatially resolved temperature–pressure profile up to an altitude of at least 400 km (refs 1, 2). In February 2012, CIRS observations showed an unexpected temperature decrease of the mesosphere by about 35 K compared to one year earlier<sup>2</sup>. Unfortunately, no CIRS limb measurements at the location of the high-altitude polar cloud exist after its appearance in May 2012. However, a set of spectra is available from 14 October 2013 that looks down on the south pole. These spectra probe temperatures at a limited range of altitudes only, but they can be used to determine whether further cooling has occurred at the south pole after February 2012. Retrievals of temperatures on 14 October 2013, using the NEMESIS



**Figure 3 | South-polar temperatures from models and retrievals.** Retrieved temperatures and their  $1\sigma$  errors at  $86^\circ\text{S}$  in September 2011<sup>1</sup> (black solid line) and February 2012<sup>2</sup> (purple) from CIRS limb measurements and at  $87^\circ\text{S}$  in October 2013 from CIRS nadir measurements (blue). We plot only regions where the observations provide reliable temperature information. Orange lines are circulation model output<sup>4</sup> for May 2012 (solid), December 2012 (dotted), and October 2013 (dashed). Black dotted lines indicate saturation temperatures for HCN volume mixing ratios of  $10^{-6}$  (left) and  $10^{-5}$  (right), which cover the measured concentrations of HCN in the south polar vortex<sup>1,2</sup>. A cloud at 300 km would require temperatures of  $\sim 125\ \text{K}$  there.

retrieval code<sup>21,22</sup>, are plotted in Fig. 3, which shows that Titan's stratosphere has cooled significantly below 200 km after 2011. Mid-infrared limb measurements of the south pole are again only planned from 2015 onwards.

The main cooling mechanism in Titan's stratosphere is thought to be radiative cooling<sup>23</sup>, and a decrease of solar irradiation at the south pole after equinox could give rise to a cold winter pole. The cooling timescale becomes shorter at greater altitudes, so it is plausible that at 300 km temperatures have dropped even more than they have below 200 km, especially since a very strong increase in trace gas concentrations has been observed since 2011<sup>1,2</sup>. These gases radiate strongly in the infrared and hence can produce a strong cooling. On the other hand, adiabatic heating due to the sinking motion of the air at the south pole has been observed in 2011 around 300 km (refs 1, 2), so the overall temperature is affected by a combination of chemistry, dynamics and insolation. The LMD global circulation model, which couples the effects of dynamics, haze formation and chemistry<sup>4,24</sup>, predicts a very warm temperature maximum around 300 km, and the temperature of this maximum is predicted to increase between 2012 and 2013 (Fig. 3). Our detection of HCN ice particles at these altitudes indicates that the polar atmosphere there is roughly 100 K colder than predicted, and thus requires the radiative cooling to be far stronger than the adiabatic heating, contrary to expectations. Hence, models of Titan's circulation require revision to understand the transitional behaviour of Titan's atmosphere around equinox.

**Note added in proof:** After acceptance of this manuscript we were made aware of near-infrared VIMS observations of Titan's northern polar hood, which also show spectral features that coincide with those of HCN ice<sup>28</sup>. Unfortunately, no further analysis was performed on that data set to confirm the presence of HCN ice by spectral modelling, or to determine its altitude.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 17 June; accepted 6 August 2014.

1. Teanby, N. A. *et al.* Active upper-atmosphere chemistry and dynamics from polar circulation reversal on Titan. *Nature* **491**, 732–735 (2012).
2. Vinatier, S. *et al.* Seasonal variations in Titan's middle atmosphere during the northern spring derived from Cassini/CIRS observation. *Icarus* (submitted).
3. West, R. A. *et al.* Post-equinox evolution of Titan's detached haze and south polar vortex cloud. Abstr. 305.03 (AAS/Division for Planetary Sciences Meeting Abstracts, Vol. 45, 2013).
4. Rannou, P., Lebonnois, S., Hourdin, F. & Luz, D. Titan atmosphere database. *Adv. Space Res.* **36**, 2194–2198 (2005).
5. Griffith, C. A., Owen, T., Miller, G. A. & Geballe, T. Transient clouds in Titan's lower atmosphere. *Nature* **395**, 575–578 (1998).
6. Samuelson, R. E., Mayo, L. A., Knuckles, M. A. & Khanna, R. J. C<sub>4</sub>N<sub>2</sub> ice in Titan's north polar stratosphere. *Planet. Space Sci.* **45**, 941–948 (1997).
7. Anderson, C. M., Samuelson, R. E., Bjoraker, G. L. & Achterberg, R. K. Particle size and abundance of HC<sub>3</sub>N ice in Titan's lower stratosphere at high northern latitudes. *Icarus* **207**, 914–922 (2010).
8. Griffith, C. A. *et al.* Evidence for a polar ethane cloud on Titan. *Science* **313**, 1620–1622 (2006).
9. Achterberg, R. K., Gierasch, P. J., Conrath, B. J., Michael Flasar, F. & Nixon, C. A. Temporal variations of Titan's middle-atmospheric temperatures from 2004 to 2009 observed by Cassini/CIRS. *Icarus* **211**, 686–698 (2011).
10. dello Russo, N. & Khanna, R. K. Laboratory infrared spectroscopic studies of crystalline nitriles with relevance to outer planetary systems. *Icarus* **123**, 366–395 (1996).
11. Moore, M. H., Ferrante, R. F., Moore, W. J. & Hudson, R. Infrared spectra and optical constants of nitrile ices relevant to Titan's atmosphere. *Astrophys. J.* **191** (suppl), 96–112 (2010).
12. Samuelson, R. E., Smith, M. D., Achterberg, R. K. & Pearl, J. C. Cassini CIRS update on stratospheric ices at Titan's winter pole. *Icarus* **189**, 63–71 (2007).
13. Lavvas, P., Griffith, C. A. & Yelle, R. V. Condensation in Titan's atmosphere at the Huygens landing site. *Icarus* **215**, 732–750 (2011).
14. Bellucci, A. *et al.* Titan solar occultation observed by Cassini/VIMS: gas absorption and constraints on aerosol composition. *Icarus* **201**, 198–216 (2009).
15. Maltagliati, L. *et al.* Titan's atmosphere as observed by VIMS/Cassini solar occultations: gaseous components. *Icarus* (submitted); preprint at <http://arxiv.org/abs/1405.6324> (2014).
16. de Kok, R., Irwin, P. G. J. & Teanby, N. A. Condensation in Titan's stratosphere during polar winter. *Icarus* **197**, 572–578 (2008).
17. Jennings, D. E. *et al.* First observation in the south of Titan's far-infrared 220 cm<sup>-1</sup> cloud. *Astrophys. J.* **761**, L15 (2012).
18. Coustenis, A., Schmitt, B., Khanna, R. K. & Trotta, F. Plausible condensates in Titan's stratosphere from Voyager infrared spectra. *Planet. Space Sci.* **47**, 1305–1329 (1999).
19. de Kok, R. *et al.* Characteristics of Titan's stratospheric aerosols and condensate clouds from Cassini CIRS far-infrared spectra. *Icarus* **191**, 223–235 (2007).
20. Anderson, C. M. & Samuelson, R. E. Titan's aerosol and stratospheric ice opacities between 18 and 500  $\mu$ m: vertical and spectral characteristics from Cassini CIRS. *Icarus* **212**, 762–778 (2011).
21. Irwin, P. G. J. *et al.* The NEMESIS planetary atmosphere radiative transfer and retrieval tool. *J. Quant. Spectrosc. Radiat. Transf.* **109**, 1136–1150 (2008).
22. Teanby, N. A., Irwin, P. G. J., de Kok, R. & Nixon, C. A. Seasonal changes in Titan's polar trace gas abundance observed by Cassini. *Astrophys. J.* **724**, L84–L89 (2010).
23. Tomasko, M. G. *et al.* Heat balance in Titan's atmosphere. *Planet. Space Sci.* **56**, 648–659 (2008).
24. Lebonnois, S., Burgalat, J., Rannou, P. & Charnay, B. Titan global climate model: a new 3-dimensional version of the IPSL Titan GCM. *Icarus* **218**, 707–722 (2012).
25. Warren, S. G. Optical constants of carbon dioxide ice. *Appl. Opt.* **25**, 2650–2674 (1986).
26. Warren, S. G. & Brandt, R. E. Optical constants of ice from the ultraviolet to the microwave: a revised compilation. *J. Geophys. Res.* **D 113**, 14220 (2008).
27. Vinatier, S. *et al.* Optical constants of Titan's stratospheric aerosols in the 70–1500 cm<sup>-1</sup> spectral range constrained by Cassini/CIRS observations. *Icarus* **219**, 5–12 (2012).
28. Clark, R. N. *et al.* Detection and mapping of hydrocarbon deposits on Titan. *J. Geophys. Res.* **115**, E10005 (2010).

**Acknowledgements** R.J.d.K. thanks the PEPS programme of the Netherlands Organisation for Scientific Research (NWO) for support. N.A.T. and P.G.J.I. were supported by the UK Science and Technology Facilities Council. L.M. thanks the Agence Nationale de la Recherche for support (ANR Project “APOSTIC” no. 11BS56002, 968 France). We thank B. Bézard, T. M. Anstey, C. Nixon and M. López-Puertas for discussions; we also thank the VIMS and CIRS operation and calibration teams.

**Author Contributions** R.J.d.K. conceived the study. R.J.d.K., L.M., N.A.T. and P.G.J.I. performed the VIMS analysis. N.A.T. and S.V. performed the CIRS analysis. All authors contributed to the interpretation, in addition to editing and improving the final manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.J.d.K. (R.J.d.Kok@sron.nl).



## METHODS

We calculated the cloud spectrum of Fig. 1b by taking the mean of the pixels containing the cloud, and combining these for 13 similar, consecutive images (v1732906961–v1732924296), weighted by their standard deviation squared. The VIMS limb spectrum is the mean from cloud pixels in Fig. 2 (image cube v1717755608). We fitted the cloud spectrum of Fig. 1b to qualitatively demonstrate the presence of HCN particles, to look for other condensates, and to obtain an estimate for the particle size. The spectrum is sensitive to the particle size, with small particles giving a stronger blue slope of the extinction cross-section, and large particles also having less pronounced absorption features. We assumed the spectrum consists of two components: a low-altitude component and a cloud component. The low-altitude component was obtained by taking the mean in two areas of roughly  $5 \times 5$  pixels on either side of the cloud, at similar solar incidence angles. Spectra from the 13 images were combined as for the cloud spectrum. The background was scaled in the fitting procedure, with the exception of the non-LTE emission at  $3.3 \mu\text{m}$ . The non-LTE emission was found to be strongly reduced in the cloud spectrum, which could be indicative of cold temperatures. The cloud component consisted of the scattering cross-section of HCN particles, multiplied by its phase function at the scattering angle of the cloud (both calculated by Mie theory using refractive indices at 120 K; ref. 11). Furthermore,

this cross-section was multiplied by the atmospheric transmission through a slant path at 250 km, as measured by VIMS<sup>15</sup>. Free parameters were the particle size, a scaling factor for the transmission spectrum, and a scaling factor for the low-altitude component, after which the cloud component was made to fit the observed spectrum between  $3.6$  and  $4.0 \mu\text{m}$ . These parameters were explored along a wide grid and for each particle size, and the best fit to the spectrum was evaluated. Particle sizes between  $0.6$  and  $1.2 \mu\text{m}$  gave qualitatively good fits to the observed spectrum. Smaller particles could not reproduce the general slope of the observed spectrum, and overestimated the  $3.2 \mu\text{m}$  feature compared to the  $4.8 \mu\text{m}$  feature. Large particles could not reproduce the spectral slope and had less pronounced absorption features in the best fit. Note that a more quantitative fit of the observations would require a three-dimensional radiative transfer model, to better model the atmospheric transmission and the low-altitude contribution near the terminator.

Temperatures were obtained from an average of  $0.5 \text{ cm}^{-1}$  resolution CIRS spectra from 14 October 2013, all within  $5^\circ$  of the south pole. We performed retrievals on the  $\nu_4$  emission band of  $\text{CH}_4$ , covering the spectra range  $1240\text{--}1360 \text{ cm}^{-1}$ , using the NEMESIS retrieval code<sup>21</sup>. We assumed a  $\text{CH}_4$  volume mixing ratio of 1.48% and previously derived haze spectral properties<sup>27</sup>. For more details on the retrieval procedure, see a previous paper<sup>22</sup>.

# Structure and evolution of the lunar Procellarum region as revealed by GRAIL gravity data

Jeffrey C. Andrews-Hanna<sup>1</sup>, Jonathan Besserer<sup>2</sup>, James W. Head III<sup>3</sup>, Carly J. A. Howett<sup>4</sup>, Walter S. Kiefer<sup>5</sup>, Paul J. Lucey<sup>6</sup>, Patrick J. McGovern<sup>5</sup>, H. Jay Melosh<sup>7</sup>, Gregory A. Neumann<sup>8</sup>, Roger J. Phillips<sup>4</sup>, Paul M. Schenk<sup>5</sup>, David E. Smith<sup>9</sup>, Sean C. Solomon<sup>10,11</sup> & Maria T. Zuber<sup>9</sup>

**The Procellarum region is a broad area on the nearside of the Moon that is characterized by low elevations<sup>1</sup>, thin crust<sup>2</sup>, and high surface concentrations of the heat-producing elements uranium, thorium, and potassium<sup>3,4</sup>. The region has been interpreted as an ancient impact basin approximately 3,200 kilometres in diameter<sup>5–7</sup>, although supporting evidence at the surface would have been largely obscured as a result of the great antiquity and poor preservation of any diagnostic features. Here we use data from the Gravity Recovery and Interior Laboratory (GRAIL) mission<sup>8</sup> to examine the subsurface structure of Procellarum. The Bouguer gravity anomalies and gravity gradients reveal a pattern of narrow linear anomalies that border Procellarum and are interpreted to be the frozen remnants of lava-filled rifts and the underlying feeder dykes that served as the magma plumbing system for much of the nearside mare volcanism. The discontinuous surface structures that were earlier interpreted as remnants of an impact basin rim are shown in GRAIL data to be a part of this continuous set of border structures in a quasi-rectangular pattern with angular intersections, contrary to the expected circular or elliptical shape of an impact basin<sup>9</sup>. The spatial pattern of magmatic-tectonic structures bounding Procellarum is consistent with their formation in response to thermal stresses produced by the differential cooling of the province relative to its surroundings, coupled with magmatic activity driven by the greater-than-average heat flux in the region.**

The Procellarum KREEP Terrane (PKT) is defined by higher than average values of the surface abundances of potassium (K), rare earth elements (REE), and phosphorus (P)<sup>3,10</sup> (Fig. 1). The PKT probably experienced a geodynamical history that differed from that of the rest of the Moon because of the elevated heat flow resulting from the high crustal concentrations of heat-producing elements<sup>10–12</sup>. The region encompasses the majority of the Moon's mare basalt provinces, including many that are not associated with known impact basins. The interpretation of the region as an impact basin was based on its distinctive composition and generally low elevation, together with the photogeological interpretation of features as fragments of circular basin rings<sup>5–7,13</sup>. The most prominent candidate ring structures are the mare shorelines and scarps on the western edge of Oceanus Procellarum and the northern edge of Mare Frigoris<sup>5</sup> (Fig. 1a; Extended Data Fig. 1). However, these arcuate segments span only a fraction of the circumference of the proposed basin, requiring much of the original topographic rim to have been later destroyed or modified beyond recognition.

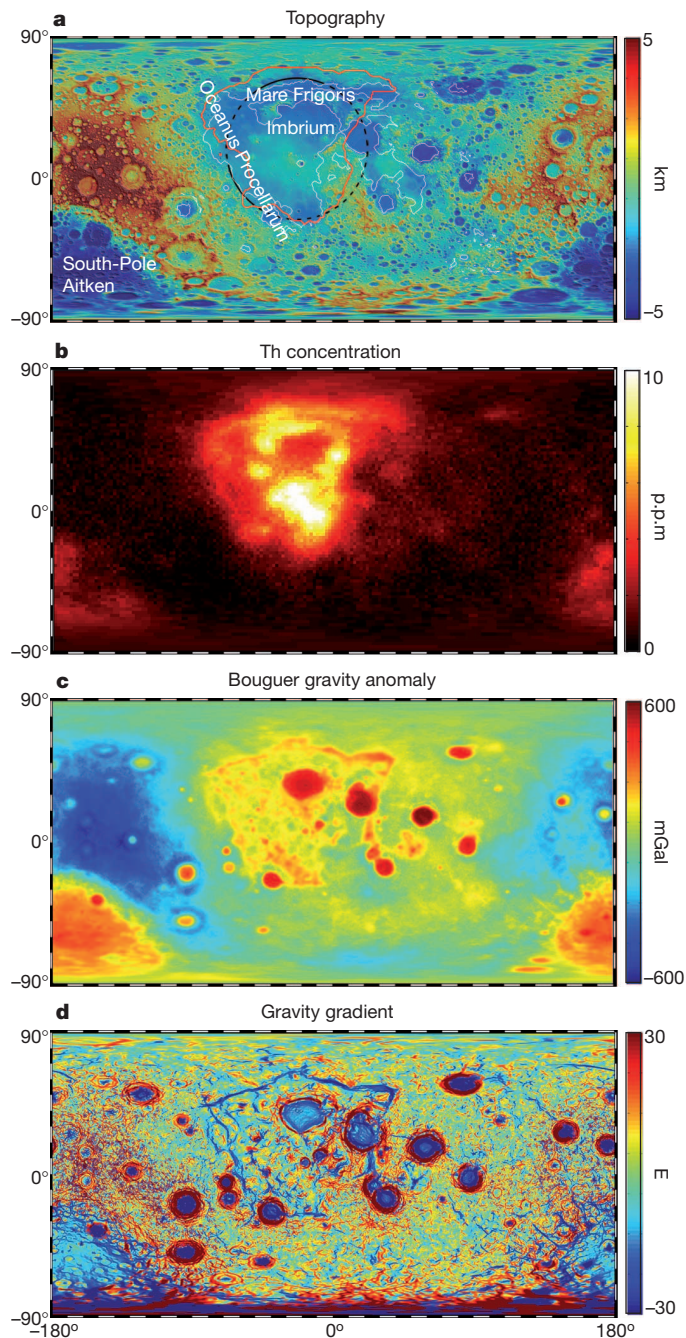
In this study, we use data from NASA's GRAIL mission<sup>8</sup> to examine the subsurface structure of the Procellarum region. Bouguer gravity anomalies (the free-air gravity field corrected for the contributions of surface topography) and gravity gradients (the second horizontal derivatives of the Bouguer potential<sup>14</sup>) reveal a distinctive pattern of anomalies surrounding the region (Fig. 1c, d). These narrow belts of negative gravity

gradients and positive gravity anomalies indicate narrow zones of positive density contrast in the subsurface. Previous analyses of the GRAIL data revealed a global population of narrow, randomly oriented, ancient igneous intrusions that lack surface expressions<sup>14</sup>. In contrast, the PKT border anomalies are broader features that are spatially associated with the maria and appear to be part of an organized large-scale structure. These anomalies are the dominant features not associated with impact basins in the global gravity gradients, but only a portion of the western border anomalies in Oceanus Procellarum were noted in earlier gravity studies<sup>15</sup>.

To investigate the source of the anomalies, we first inverted the gravity field in the spherical harmonic domain under the assumption that the anomalies arise from variations in the thickness of both the maria and the underlying feldspathic crust that serves as the basement of the maria (see Methods for details). We focus here on two models to illustrate the range of solutions: the first imposes an isostatic condition on the pre-mare crust, and the second forces the amplitude of the relief along the mare–basement and crust–mantle interfaces to be equal and opposite in magnitude. For these two models, the average structures across two of the border anomalies at the northwest corner of the PKT suggest the presence of elongated mare-filled depressions in the feldspathic crust having widths of ~150 km and depths of 2–4 km, and underlain by crust–mantle interfaces that are shallower than adjacent areas by 3–6 km (Fig. 2e–h; Extended Data Figs 2, 3). If we instead assume that the PKT border anomalies arise from igneous intrusions in the subsurface<sup>14</sup>, inversions of the average gravity profiles across these two anomalies yield widths of  $66^{+5}_{-6}$  and  $82^{+19}_{-36}$  km and vertical extents of  $8^{+1}_{-1}$  and  $6^{+3}_{-1}$  km for intrusions with elliptical cross-sections, assumed density contrasts of  $550 \text{ kg m}^{-3}$ , and bottom depths of 25 km (Fig. 2c, d; see Methods).

The spherical harmonic inversion solutions are consistent with thickening of the maria over linear depressions formed by crustal thinning, as could occur in volcanically flooded rift valleys<sup>16</sup>. The branching of anomalies that make up the western border structure and the triple-junction intersections at some corners are consistent with the attributes of planetary rifts. This interpretation is also supported by the broad elongated depressions surrounding the border anomalies beneath Mare Frigoris and western Mare Tranquillitatis, and the scarps found in the highlands adjacent to some of the border anomalies<sup>5</sup>. The inferred crustal thinning could arise from extension of the crust by 8–18 km (Extended Data Table 1). For the intrusion models, the large widths of the inferred intrusions (greatly exceeding the vertical dimensions), and the association of the gravity anomalies with maria at the surface, suggest that dyke-like intrusions are not solely responsible for the anomalies. A combination of crustal thinning, mare thickening, and intrusion by dyke swarms provides the most likely explanation for the anomalies. The

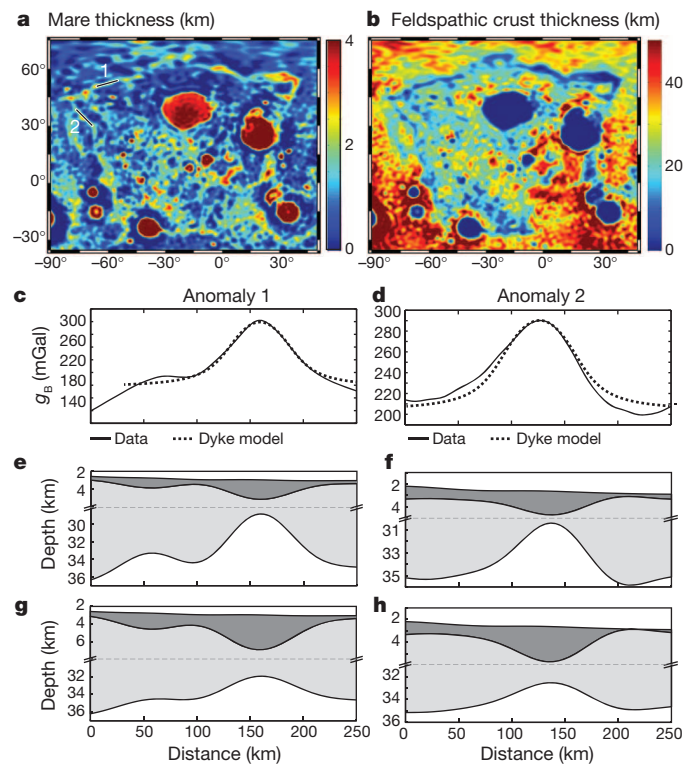
<sup>1</sup>Department of Geophysics and Center for Space Resources, Colorado School of Mines, Golden, Colorado 80401, USA. <sup>2</sup>Department of Earth and Planetary Sciences, University of California, Santa Cruz, California 95064, USA. <sup>3</sup>Department of Earth, Environmental and Planetary Sciences, Brown University, Providence, Rhode Island 02912, USA. <sup>4</sup>Planetary Science Directorate, Southwest Research Institute, Boulder, Colorado 80302, USA. <sup>5</sup>Lunar and Planetary Institute, Houston, Texas 77058, USA. <sup>6</sup>Hawaii Institute of Geophysics and Planetology, University of Hawaii, Honolulu, Hawaii 96822, USA. <sup>7</sup>Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, Indiana 47907, USA. <sup>8</sup>Solar System Exploration Division, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA. <sup>9</sup>Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, USA. <sup>10</sup>Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington DC 20015, USA. <sup>11</sup>Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York 10964, USA.



**Figure 1 | Global maps of lunar properties.** **a**, Topography; **b**, Th concentration; **c**, Bouguer gravity anomaly; and **d**, gravity gradient (in units of Eötvös;  $1 \text{ E} = 10^{-9} \text{ s}^{-2}$ ). All maps are simple cylindrical projections centred on the nearside. The circular rim of the proposed Procellarum impact basin<sup>5</sup> (black dashed line), the outline of the maria (white lines<sup>17</sup>), and the extent of the PKT (red line, corresponding to a Th concentration of 3.5 p.p.m.; ref. 4) are shown in **a**. Features discussed in the text are labelled in **a**.

elevated heat flux in the PKT<sup>10</sup> coupled with passive mantle upwelling during rifting would have led to widespread partial melting of the underlying mantle<sup>16</sup>, so tectonic extension would have been accompanied by dyke intrusion and volcanism. These dykes may represent the magma plumbing system that provided conduits connecting deep magma reservoirs to many of the nearside maria.

The PKT border structures are the only known lunar structures consistent with large-scale rifting of the crust, a process that is more common on Earth, Venus, and Mars. The surface exposures of the maria overlying the border structures formed  $3.51 \pm 0.25$  billion years ago

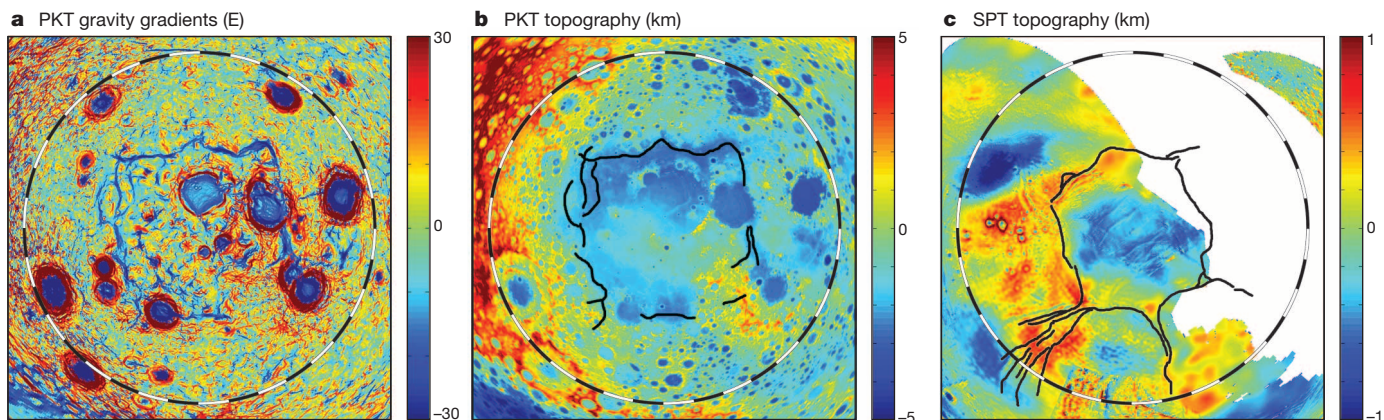


**Figure 2 | Gravity and subsurface structure of the PKT border structures.** **a**, **b**, Maps of the modelled thickness of the maria (**a**) and underlying feldspathic crust (**b**) assuming that the mare–basement and crust–mantle interfaces were in isostatic equilibrium before infilling by mare basalt. **c**, **d**, Profiles of the average Bouguer gravity anomaly  $g_B$  perpendicular to border anomalies 1 (**c**) and 2 (**d**); see **a** for locations. The dashed lines show the predicted gravity for the best-fit dykes. **e**–**h**, Average cross-sections of the model results orthogonal to border anomalies 1 (**e**, **g**) and 2 (**f**, **h**) showing the mare (dark grey) and feldspathic crust (light grey) for two different sets of filters. The filters used in the models in **e** and **f** impose the isostatic condition as in **a** and **b**, whereas the filters used in the models in **g** and **h** impose the condition that the relief along the interfaces was equal and opposite in amplitude (see Methods for further details and Extended Data Fig. 3 for results from additional models).

(Gyr ago; area-weighted mean and standard deviation)<sup>17</sup>, representing the final stages of the volcanic infilling of the structures. In contrast, the rest of the nearside maria exhibit a range of surface ages of 1.2–4.0 Gyr. Volcanic infilling of the rifts may have been a self-limiting process because the flexural response to the loading would have caused compression in the upper lithosphere, possibly closing off the magma conduits. This inference is supported by the observation of wrinkle ridges overlying and parallel to the border structures. Parallel wrinkle ridges flanking the Mare Frigoris border structure may also reflect structural control of the wrinkle ridges by buried tectonic structures.

In a polar projection centred on the PKT, the border structures delineate a quasi-rectangular shape  $\sim 2,600$  km in width (Fig. 3). The arcuate scarps at the edges of Mare Frigoris and Oceanus Procellarum that were previously interpreted as rim segments of a Procellarum basin are seen in the GRAIL data to be a small fraction of this continuous set of well-expressed structures that trace out a polygonal pattern consisting of predominantly straight sides and angular intersections (Extended Data Fig. 1). The northeast and northwest corners of the structure deviate from the proposed circular rim<sup>5</sup> by  $\sim 215$  km and  $\sim 175$  km, respectively. Only the discontinuous and poorly expressed anomalies in the southwestern portion of the region are compatible with a circular rim. This quasi-rectangular pattern is in contrast with the circular or elliptical shapes of all other large impact basins<sup>9</sup>, including the ancient hemisphere-scale Borealis basin on Mars, for which a continuous elliptical basin rim can be traced in topography and gravity data<sup>18</sup>. The interpretation of





**Figure 3 | Geometric pattern of the PKT border structures, with a comparison to the Enceladus SPT.** **a, b,** The border structures of the PKT highlighted by the gravity gradients **(a)** trace out a quasi-rectangular pattern, enclosing a broad region of low elevations<sup>1</sup> **(b)**. **c,** The SPT is similarly a region of low elevation<sup>25</sup> (white regions denote topography data gaps) and high

heat flow<sup>26</sup> (Extended Data Fig. 8) surrounded by a quasi-rectangular pattern of border structures. The black lines in **b** and **c** trace the border structures surrounding the PKT and SPT, respectively. All maps are in a simple polar projection; in all panels, the circle corresponds to an angular diameter of 180° of surface arc, divided into 10° increments.

the PKT border structures as the rim of an impact basin would require hundreds of kilometres of horizontal deformation with large strain gradients to produce the angular corners, but there is no evidence for such large-magnitude strain on the Moon<sup>19</sup>. Furthermore, the negative gravity gradients of the border structures do not match the signatures of known impact basins, such as the Imbrium and South Pole–Aitken basins, which are characterized by paired positive and negative gradients of equal amplitude flanking the rims and negative gradients throughout the basin interiors. Although it is not possible to disprove the existence of an ancient degraded Procellarum basin that lacks a clear geophysical signature, the geometry and gravitational signature of the structures bordering the PKT do not support the interpretation that they mark the rim of a basin.

The formation and geometric pattern of the PKT border structures require an explanation. Although the gravity anomalies are consistent with either lava-flooded rift valleys or dense swarms of dykes, both interpretations require substantial extension across the border structures. The locations of the structures at the edge of the PKT suggest that the elevated heat flux in this region<sup>10</sup> may have played a role in the extension inferred from the gravity modelling. In a state of thermal equilibrium, both the temperature and the rate of change in temperature at a given depth in the lithosphere would be linearly proportional to the concentration of heat-producing elements in and/or beneath the crust. Thus, although the PKT was always warmer than its surroundings owing to the high concentrations of heat-producing elements, it would have cooled at a greater rate as a result of the declining radiogenic heat production<sup>10</sup>. The cooling lithosphere would then have experienced thermal contraction, which in turn would have caused horizontal extension at the margins. Cooling by 600 K across a region 2,000 km wide would have induced the equivalent of ~8 km of extension. We tested this hypothesis with a simple model of the thermal evolution and resultant stresses (see Methods). A finite difference model was used to represent the conductive thermal evolution of the Moon, given the equivalent of 10 km of KREEP basalt at the base of a 40-km-thick crust within a spherical cap 2,000 km in diameter<sup>10,11</sup>. The model predicts a temperature decrease beneath the PKT relative to its surroundings of as much as 600 K between 4.0 and 3.0 Gyr ago, with the maximum cooling at the base of the crust (Fig. 4a; Extended Data Figs 4, 5).

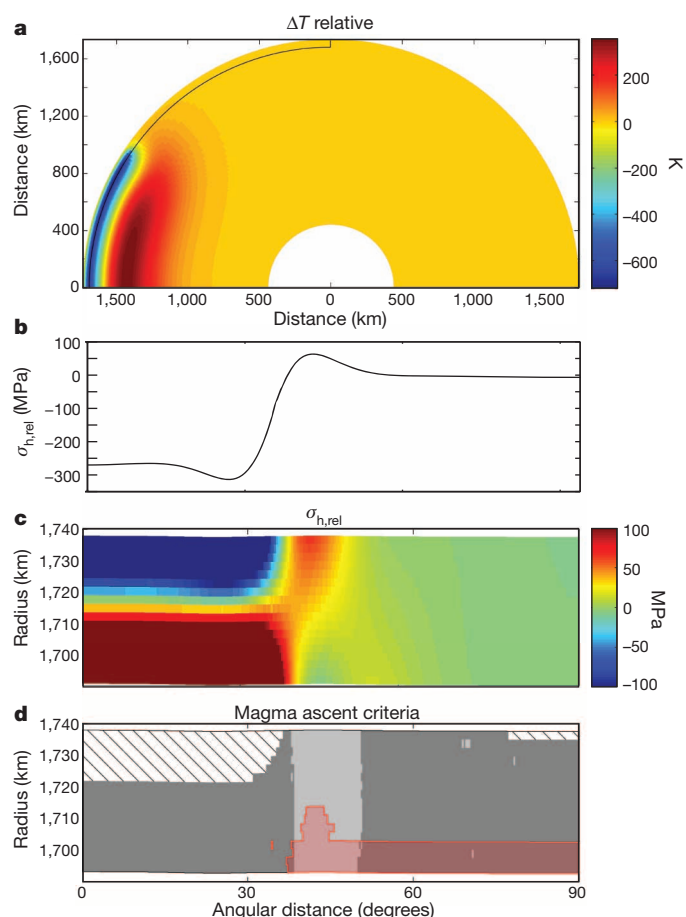
The stresses resulting from the thermal contraction of the lithosphere between 4.0 and 3.0 Gyr ago were calculated with an elastic finite element model<sup>20</sup>. The far-field stresses on the opposite side of the Moon were subtracted in order to isolate the effects of the PKT, because the mean stress in the lithosphere may have been affected by global contraction or expansion<sup>14,21</sup>. Cooling and contraction of the lower lithosphere

within the PKT caused extension, which induced compression in the elastically coupled upper lithosphere inside the PKT, and extension throughout the lithosphere at the edge of the PKT (Fig. 4c). Similar results were obtained if the KREEP-rich material was distributed throughout the crust (Extended Data Figs 6, 7). This extension may have been augmented by an early period of global expansion<sup>14</sup>.

Many of the maria not associated with unambiguous impact basins are found over the PKT border structures, including maria Nubium, Procellarum, Frigoris, Mortis, Somniorum, and Tranquillitatis. Rise of magma to the surface in dykes requires that the greatest tensile stress be horizontal, and a vertical gradient in stress that is conducive to magma ascent<sup>22</sup>. The model predicts that the extensional zone bordering the PKT was conducive to magma ascent in dykes (Fig. 4d). In contrast, horizontal compressional stresses in the upper lithosphere within the centre of the PKT would tend to inhibit the rise of magma, except where this stress field was modified by later processes such as impacts or loading and flexure of the lithosphere, or where magma ascent was aided by volatile exsolution or a pressurized magma chamber.

In order to form the observed rectilinear pattern of structures, it is necessary to break the azimuthal symmetry assumed in the model. Volumetric contraction beneath a free surface generates fracture patterns with characteristic corner angles of 120°. This pattern results in six-sided polygons at scales ranging from 1–100 cm (for example, mud cracks, columnar joints in basalt), to 1–100 m (for example, thermal contraction polygons in permafrost), to 10 km (for example, polygons from sediment compaction in the lowlands of Mars<sup>23</sup>). However, as the size of the structure becomes large relative to the radius of the planetary body, surface curvature becomes important. A polygon with 120° corner angles will have five or four sides when the lengths of the sides reach 41.8° or 70.5° of arc, respectively. The mean length of the PKT border structures is 2,150 km or 71°, and the angles of the vertices range from 109° to 125°. Thus, at the scale of the PKT, a set of linear rifts intersecting at 120°-angle junctions around a contracting cap may result in a quasi-rectangular structure.

We note a similarity in the pattern of structures to the south polar terrain (SPT) of Saturn's icy moon Enceladus (Fig. 3; Extended Data Fig. 8)<sup>24,25</sup>. Both the PKT and SPT are bordered by quasi-rectangular sets of tectonic belts with angular intersections that sometimes take the form of triple junctions. Both structures enclose regions approximately 70–80° in diameter of low topography<sup>1,25</sup>, enhanced volcanic activity<sup>10,24</sup>, and strongly elevated heat flow<sup>10,26</sup>. However, we emphasize that there are important differences between the specific processes at work and the evolutionary histories of these two different terrains, including (on Enceladus) the tidal source of the heat, the prevalence of compressional



**Figure 4 | Predicted temperature and stress for the Procellarum region.**

**a**, Predicted temperature change of the PKT relative to its surroundings between 4.0 and 3.0 Gyr ago. The Procellarum region is centred on the pole on the left side of the figure. The black line denotes the area expanded in **c** and **d**. **b**, In-plane horizontal elastic relative stress radial to the centre of the PKT at the surface predicted by the finite element model (where positive stresses are tensile; the far-field stress profile has been subtracted to calculate the relative stresses). **c**, Cross-section of the in-plane horizontal elastic relative stress. **d**, Predicted zones of magma ascent; dark grey indicates horizontal extension conducive to vertical dyke formation, light grey indicates both horizontal extension and a vertical stress gradient more favourable to magma ascent than in the lithosphere far from the PKT, and red indicates areas in which magma will rise unassisted by other factors. Cross-hatching indicates regions in which none of the criteria for magma ascent are met. The temperatures in **a** and stresses in **b**, **c** are both taken relative to the far-field values in the opposite hemisphere.

tectonics<sup>24,25</sup>, the likelihood of a subsurface ocean<sup>27</sup>, and the possibility of a mobile lithosphere<sup>28</sup>. Nevertheless, the gross morphological and geophysical similarities between the PKT on the Moon and the SPT on Enceladus suggest the possibility of broad parallels in their geodynamic evolution, and that similar parallels may exist with other magmatic-tectonic centres (for example, the northern lowlands of Mercury, an irregular depression  $\sim 80^\circ$  in diameter<sup>29</sup> that has experienced widespread volcanic resurfacing<sup>30</sup>).

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 18 January; accepted 16 July 2014.**

- Smith, D. E. *et al.* Initial observations from the Lunar Orbiter Laser Altimeter (LOLA). *Geophys. Res. Lett.* **37**, L18204 (2010).

- Wieczorek, M. A. *et al.* The crust of the Moon as seen by GRAIL. *Science* **339**, 671–675 (2013).
- Jolliff, B. L. *et al.* Major lunar crustal terranes: surface expressions and crust-mantle origins. *J. Geophys. Res.* **105**, 4197–4216 (2000).
- Lawrence, D. J. *et al.* Global elemental maps of the Moon: the Lunar Prospector gamma-ray spectrometer. *Science* **281**, 1484–1489 (1998).
- Whitaker, E. A. The lunar Procellarum basin. *Lunar Planet. Sci.* **12A**, 105–111 (1981).
- Cadogan, P. H. Oldest and largest lunar basin? *Nature* **250**, 315–316 (1974).
- Wilhelms, D. E. The geological history of the Moon. *US Geol. Surv. Prof. Pap.* **1348**, 1–302 (1987).
- Zuber, M. T. *et al.* Gravity field of the Moon from the Gravity Recovery and Interior Laboratory (GRAIL) mission. *Science* **339**, 668–671 (2013).
- Andrews-Hanna, J. C. & Zuber, M. T. in *Large Meteorite Impacts and Planetary Evolution IV* (eds Gibson, R. L. & Reimold, W. U.) 1–13 (Special Paper 465, Geological Society of America, 2010).
- Wieczorek, M. A. & Phillips, R. J. The “Procellarum KREEP Terrane”: implications for mare volcanism and lunar evolution. *J. Geophys. Res.* **105**, 20417–20430 (2000).
- Grimm, R. E. Geophysical constraints on the lunar Procellarum KREEP Terrane. *J. Geophys. Res. Planets* **118**, 768–777 (2013).
- Laneville, M., Wieczorek, M. A., Breuer, D. & Tosi, N. Asymmetric thermal evolution of the Moon. *J. Geophys. Res. Planets* **118**, 1435–1452 (2013).
- Nakamura, R. *et al.* Compositional evidence for an impact origin of the Moon’s Procellarum basin. *Nature Geosci.* **5**, 775–778 (2012).
- Andrews-Hanna, J. C. *et al.* Ancient igneous intrusions and early expansion of the Moon revealed by GRAIL gravity gradiometry. *Science* **339**, 675–678 (2013).
- Scott, D. H. The geologic significance of some lunar gravity anomalies. *Proc. Lunar Sci. Conf.* **5**, 3025–3036 (1974).
- White, R. & McKenzie, D. Magmatism at rift zones: the generation of volcanic continental margins and flood basalts. *J. Geophys. Res.* **94**, 7685–7729 (1989).
- Hiesinger, H. *et al.* Ages and stratigraphy of lunar mare basalts in Mare Frigoris and other nearside maria based on crater size-frequency distribution measurements. *J. Geophys. Res.* **115**, E03003 (2010).
- Andrews-Hanna, J. C., Zuber, M. T. & Banerdt, W. B. The Borealis basin and the origin of the martian crustal dichotomy. *Nature* **453**, 1212–1215 (2008).
- Watters, T. R. & Johnson, C. L. in *Planetary Tectonics* (eds Watters, T. R. & Schultz, R. A.) 121–182 (Cambridge Univ. Press, 2010).
- Melosh, H. J. & Raefsky, A. The dynamical origin of subduction zone topography. *Geophys. J. R. Astron. Soc.* **60**, 333–354 (1980).
- Solomon, S. C. The relationship between crustal tectonics and internal evolution in the Moon and Mercury. *Phys. Earth Planet. Inter.* **15**, 135–145 (1977).
- McGovern, P. J., Rumpf, M. E. & Zimbelman, J. R. The influence of lithospheric flexure and volcano shape on magma ascent at large volcanoes on Venus. *J. Geophys. Res. Planets* **118**, 2423–2437 (2013).
- Hiesinger, H. & Head, J. W. Characteristics and origin of polygonal terrain in southern Utopia Planitia, Mars: results from Mars Orbiter Laser Altimeter and Mars Orbiter Camera data. *J. Geophys. Res.* **105**, 11999–12022 (2000).
- Porco, C. C. *et al.* Cassini observed the active south pole of Enceladus. *Science* **311**, 1393–1401 (2006).
- Schenk, P. M. & McKinnon, W. B. One-hundred-km-scale basins on Enceladus: evidence for an active ice shell. *Geophys. Res. Lett.* **36**, L16202 (2009).
- Howett, C. J. A., Spencer, J. R., Pearl, J. & Segura, M. High heat flow from Enceladus’ south polar region measured using 10–600 cm<sup>−1</sup> Cassini/CIRS data. *J. Geophys. Res.* **116**, E03003 (2011).
- Běhouňková, M., Tobie, G., Choblet, G. & Cadek, O. Tidally-induced melting events as the origin of south-pole activity on Enceladus. *Icarus* **219**, 655–664 (2012).
- O’Neil, C. & Nimmo, F. The role of episodic overturn in generating the surface geology and heat flow on Enceladus. *Nature Geosci.* **3**, 88–91 (2010).
- Zuber, M. T. *et al.* Topography of the northern hemisphere of Mercury from MESSENGER laser altimetry. *Science* **336**, 217–220 (2012).
- Head, J. W. *et al.* Flood volcanism in the northern high latitudes of Mercury revealed by MESSENGER. *Science* **333**, 1853–1856 (2011).

**Acknowledgements** The GRAIL mission is a component of the NASA Discovery Program and is performed under contract to the Massachusetts Institute of Technology and the Jet Propulsion Laboratory, California Institute of Technology. J.C.A.-H. was supported by grant NNX12AL20G from NASA’s GRAIL Guest Scientist Program.

**Author Contributions** J.C.A.-H. performed the data analyses and modelling. M.T.Z. is the principal investigator of the GRAIL mission. All authors contributed to the interpretation of the results and their implications.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.C.A.-H. ([jcahanna@mines.edu](mailto:jcahanna@mines.edu)).



## METHODS

**Gravity gradients.** The gravity data analysed here were taken from gravity model GRGM900b, obtained from observations during GRAIL's primary and extended missions<sup>31</sup>. The Bouguer gravity anomaly model was generated for an assumed a crustal density of  $2,550 \text{ kg m}^{-3}$  (ref. 2). The Bouguer gravity gradients were calculated in the spherical harmonic domain<sup>32</sup> using the software archive SHTOOLS (freely available on-line at <http://shtools.ipgp.fr>). The eigenvalues of the horizontal gravity gradient tensor ( $\Gamma_{11}$ ,  $\Gamma_{22}$ ), representing the values of the maximum and minimum curvature of the potential field at each point, were then calculated. As was done previously<sup>14</sup>, the eigenvalues were combined into a single value (the maximum-amplitude horizontal gradient, or  $\Gamma_{hh}$ ) representing the second horizontal derivative of maximum amplitude at each point on the surface:

$$\Gamma_{hh} = \begin{cases} \Gamma_{11} & \text{if } |\Gamma_{11}| > |\Gamma_{22}| \\ \Gamma_{22} & \text{if } |\Gamma_{11}| \leq |\Gamma_{22}| \end{cases}$$

where  $|x|$  indicates the absolute value of  $x$ . This maximum-amplitude horizontal gradient represents the gradient orthogonal to any structures that dominate the local gravity, regardless of their orientation. The gravity gradients are given in units of Eötvös ( $1 \text{ E} = 10^{-9} \text{ s}^{-2}$ ). The gravity gradients were used to reveal the presence of discrete subsurface structures, whereas the Bouguer gravity anomaly and potential were used in all subsequent analyses.

In this representation of the gravity gradients, a positive density anomaly will produce a negative gravity gradient, whereas a step function density anomaly will produce a symmetric pair of positive and negative gravity gradients flanking the step. For this reason, the mantle uplift beneath large impact basins is expressed as an outer ring of positive gravity gradients and an inner ring of negative gravity gradients. Thus, although some of the border structures are near the edges of the overlying maria, the gravity gradient signatures are not consistent with the anomalies expected to arise from edge effects of the maria. Furthermore, the northern border anomaly is approximately centred within Mare Frigoris, and the western border structure exhibits three branches that are offset from the edge of the overlying Oceanus Procellarum by as much as 600 km. The average Bouguer gravity profiles perpendicular to the border structures reveal narrow positive Bouguer anomalies (Fig. 2c, d). The elongated negative gravity gradients and positive Bouguer gravity anomalies bordering the Procellarum KREEP Terrane (PKT) are most simply explained by elongated positive density anomalies.

In previous work focusing on narrower structures in the lunar gravity gradient field interpreted as elongated igneous intrusions or swarms of dykes, we calculated the gradients using a high-pass filter at degree and order 50, which emphasized shorter-wavelength structures<sup>14</sup>. The focus of the present work is on the longer-wavelength border anomalies surrounding the PKT, which have significant power at degrees less than 50. Thus, the gravity gradients were calculated between degrees 2 and 400, with a cosine-shaped taper applied between degrees 350 and 400. Two of the border anomalies in the northwest part of the region coincide with ancient igneous intrusions identified in the previous study of the short-wavelength gravity gradients<sup>14</sup>. However, the majority of the dyke-like structures identified in that study are narrower features that lack a surface expression and appear to be distributed randomly across the planet<sup>14</sup>. In contrast, the PKT border anomalies are longer-wavelength features that occur within the maria and appear to be part of a large-scale organized structure.

In order to highlight the true shape of the PKT border anomalies, the Bouguer gravity gradients were plotted in a simple polar projection, preserving the distance between each point and the origin, and thus preserving the shape of features centred on the origin. The global Bouguer gravity gradient map in cylindrical projection (Fig. 1) appears to show a pentagonal structure encompassing the PKT. However, re-projection in a polar projection centred on the region (Fig. 3a) reveals that the structure as a whole is dominantly quasi-rectangular. The pentagonal appearance in the cylindrical projection is a result of both the distortions at high latitudes in that projection and a kink in the northern border structure at its mid-point.

A previous study<sup>7</sup> mapped possible ring structures associated with the Procellarum basin on a Lambert azimuthal equal-area map of the nearside of the Moon. A comparison of the GRAIL gravity gradients with this map (Extended Data Fig. 1) reveals that the majority of the mare shorelines and major scarps identified in that study parallel the Procellarum border anomalies, and a substantial fraction of the wrinkle ridges overlie the border anomalies. However, the angular corners apparent in the gravity gradients are missing or rounded off in the mapped surface structures. The scarps and mare shorelines adjacent to the border anomalies are consistent with their interpretation as lava-flooded rifts, and the alignment of wrinkle ridges over the border anomalies is consistent with the flexural stresses expected to arise from the narrow loads inferred from the gravity data. The tracing of these structures on a Lambert azimuthal equal-area map, which does not preserve angles and causes significant distortions around the edges due to the nonlinear radial distance

scale, contributes to the apparent circularity of the structures. This distortion is particularly prominent for the northwest corner of the PKT border structures, which occurs near the limb of the Moon where the distortion is at its greatest. Nevertheless, even in this projection the border anomalies clearly delineate a polygonal structure. A simple polar projection centred on the Procellarum region preserves the distance from the centre to all points and thus provides a more accurate depiction of shapes centred on the origin. Only the discontinuous structures in the southwest corner of the Procellarum region are consistent with a circular pattern.

**Gravity inversions.** Long-wavelength Bouguer gravity anomalies on the Moon are thought to arise largely from variations in the relief along the crust–mantle interface<sup>2,33</sup>. In contrast, because the gravitational potential of short-wavelength anomalies attenuates rapidly with elevation, most of the observed high-degree power in the Bouguer gravity must arise from density variations at depths shallower than the crust–mantle interface. At intermediate degrees, the origin of the gravity anomalies depends on the geodynamic setting. For the case of the PKT, the vast majority of the border anomalies occur beneath maria, and thus the anomalies probably arise at least in part from variations in the relief along the mare–basement interface. However, some minor branches extend off from the main border anomalies into the surrounding crust outside the maria, suggesting that at least some component of intrusive dykes and/or uplifted crust–mantle interface contributes to the anomalies. We consider both possibilities in our analysis.

The width of the gravity anomalies and their association with mare basalts at the surface suggest that the anomalies may be the result of local thickening of the maria above linear tectonic structures and/or uplift of the crust–mantle interface beneath those structures. To investigate this scenario, we inverted the gravity data in the spherical harmonic domain by downward continuing the Bouguer gravity to the appropriate radii and iteratively solving for the spherical harmonic coefficients describing the relief along the density interfaces of interest, taking into account the finite-amplitude effects of that relief<sup>33</sup>. This approach has been applied previously for calculating the relief along the crust–mantle interface<sup>2,33</sup>, but here we wish to solve for the relief along both the mare–basement and crust–mantle interfaces. We first calculated the Bouguer gravity anomaly field using the density of mare basalt, since the maria comprise the top layer in our three-layer model (mare, feldspathic crust, and mantle). We adopt a mare density of  $\rho_m = 3,150 \text{ kg m}^{-3}$ , based on the average of measured densities of Apollo mare samples<sup>34</sup>. The Bouguer anomaly was then used to calculate the relief along the mare–basement and crust–mantle interfaces.

The solution for the relief along two different subsurface density interfaces is inherently non-unique. In order to capture a range of possible solutions, we consider different filters to parse the gravity anomalies between the crust–mantle interface and the mare–basement interface. We designed a filter  $w_l$  to allow us to specify the desired ratio,  $f$ , between the relief along the crust–mantle interface and that along the mare–basement interface, taking into account the degree-dependent amplification of the gravity anomalies during their downward continuation to the mean depth of the interface of interest:

$$w_l = \frac{(R_m/R_0)^{l+2}(\rho_m - \rho_c)f}{(R_m/R_0)^{l+2}(\rho_m - \rho_c) + (R_m/R_0)^{l+2}(\rho_m - \rho_c)f}$$

where  $l$  is the spherical harmonic degree,  $\rho_c$  is the density of the feldspathic crust,  $\rho_m$  is the density of the mantle,  $\rho_m$  is the density of mare material,  $R_0$  is the mean lunar radius (1,737.15 km; ref. 1),  $R_m$  is the mean radius of the mare–basement interface, and  $R_m$  is the mean radius of the crust–mantle interface. This filter was applied in calculating the relief along the crust–mantle interface, and the remaining Bouguer gravity was then used to calculate the relief along the mare–basement interface. We assumed densities of  $2,550 \text{ kg m}^{-3}$  and  $3,220 \text{ kg m}^{-3}$  for the porous feldspathic crust and mantle, respectively, on the basis of previous GRAIL analyses<sup>2</sup>. We assumed a mean radius of the crust–mantle interface of 1,697.15 km, resulting in a mean crustal thickness of 40 km, and a mean radius of the mare–basement interface of 1,736.15 km. The filters used for the models depicted in Fig. 2 are shown in Extended Data Fig. 2. The first model represents the case in which the feldspathic crust was in a state of isostasy before infilling by the mare, leading to a ratio  $f$  of  $\rho_c/(\rho_m - \rho_c)$  (Extended Data Fig. 2a). In this model, isostasy is defined using the simple criterion of equal masses in adjacent columns. If some of the volcanic infilling of the structures occurred in parallel with the extensional tectonics, the resulting load would have driven added subsidence, which would have increased the ratio between the relief along the mare–basement and crust–mantle interfaces. The second model represents the case in which the relief along the two interfaces was equal and opposite in amplitude, with  $f$  taking on a value of 1 for degrees  $> 10$ . However, because the long-wavelength topography of the Moon is largely isostatic, we adopted the isostatic ratio for  $f$  for degrees 1–3, with a linear transition between the isostatic and equal-amplitude values over degrees 3–10, and the equal-amplitude value from degrees 10 to 125 (Extended Data Fig. 2b). These two models serve to illustrate the range of possible solutions and the relative insensitivity of the inferred



extension to the model assumptions. A low-pass cosine taper from degrees 125 to 150 was applied to all models.

The resulting models match the gravity data but do not take into account the effects of flexure, which would perturb the interface depths relative to their elevations before mare loading and thus alter the assumed pre-loading ratio between the relief along the interfaces. Although the models were applied globally, the results are not valid in areas outside the maria. Similarly, crustal thickness models that neglect the high density of the mare basalt and the possible variations in mare thickness will have errors within the maria. The mean radius of the mare–basement interface was chosen so as to bring the base of the maria within the Procellarum region below the surface over most of the observed maria. However, the modelled long-wavelength variations in the thickness of the maria are poorly constrained because of the ambiguity between the gravitational effects of variations in the relief along the mare–basement and crust–mantle interfaces. As a result, the distribution of areas with predicted mare thicknesses greater than zero only approximately matches the observed distribution of the maria. Nevertheless, the short-wavelength variations in the thickness of the maria beneath the border anomalies are robust, given the model assumptions.

The density of the lunar mantle beneath the PKT is not known. The process responsible for concentrating the KREEP-rich materials on the nearside of the Moon may have also brought dense ilmenite-rich cumulates to the base of the crust on the nearside<sup>35</sup>. Overturn of the buoyantly unstable magma ocean cumulates would have mixed this material to deeper levels in the lunar mantle<sup>36,37</sup>, but this overturn may have been limited by the high viscosity of the solid ilmenite-rich cumulates and is predicted to have occurred only for a limited range of scenarios<sup>38</sup>. It is possible that a mixture of olivine and ilmenite-rich cumulates sank as solid diapirs, leaving behind a portion of the ilmenite-rich material at shallower levels<sup>38</sup>. To account for the possibility of shallow ilmenite-rich material beneath the PKT, we considered a high-mantle-density end-member model with an assumed mantle density of  $3,500 \text{ kg m}^{-3}$ , representative of the density of the late-stage crystallization products from the magma ocean<sup>37</sup>. The higher mantle density reduces the predicted mantle uplift beneath the border structures, and similarly reduces the predicted extension.

We also considered two additional end-member scenarios in our gravity models. For one model, we assumed that all of the gravity anomalies at degrees  $>10$  arise from variations in the thickness of the maria. This model required a mean mare–basement interface radius of  $R_0 - 6 \text{ km}$  in order to bring the mare–basement interface below the surface in the regions of interest. For another model, we assumed that all of the gravity anomalies at degrees  $>10$  arise from variations in the relief along the crust–mantle interface. This model became unstable at higher degrees because of the amplification of the high-degree gravity anomalies during downward continuation to the mean depth of the crust–mantle interface, so a cosine taper was applied between degrees 75 and 100 to stabilize the solution. As a result, this model is a factor of 1.6 coarser in resolution than the other models. This result provides further evidence that the short-wavelength gravity anomalies must arise from density anomalies at depths more shallow than the crust–mantle boundary. This model ascribing all of the Bouguer gravity anomaly field to variations along the crust–mantle interface is comparable in resolution to the global GRAIL crustal thickness models<sup>2</sup> (low-pass filtered with an amplitude of 0.5 at degrees 87 and 80, respectively, corresponding to spatial wavelengths of 63 and 68 km). In contrast, the models ascribing a substantial fraction of the Bouguer gravity field to the shallower mare–basement interface are higher in resolution (low-pass filtered with an amplitude of 0.5 at degree 137, corresponding to a spatial wavelength of 40 km). For both of these models, we assumed that variations in the top and bottom surfaces of the feldspathic crust from degrees 1 to 3 were isostatically compensated before mare flooding, with a linear transition to the desired filter from degrees 3 to 10. These final two models are not likely to be accurate representations of the subsurface structure, but they bracket the range of possible solutions.

The predicted relief along the interfaces was used to calculate the thicknesses of the feldspathic crust and maria (Extended Data Fig. 3). The broad patterns of mare thickness in this region as indicated by the models are highly uncertain because of the non-uniqueness of the division of the gravity anomalies between the mare–basement and the crust–mantle interfaces. In some areas, the predicted base of the mare rises above the surface, indicating the need for subsurface mass deficits such as those that could arise from additional variations in the crustal thickness or density in order to explain the observed gravity field within the context of this model. These errors outside the maria do not affect the predictions for the Procellarum border structures. The local thickening of the mare over the western Procellarum border structure is broadly consistent with maps of the mare thickness derived from geological constraints, such as the burial depths of impact craters, which show local thickenings of up to  $>1.5 \text{ km}$  along this structure<sup>39</sup>. Models combining the effects of dykes with the relief along the mare–basement and crust–mantle interfaces would predict narrower dykes than models that ascribe the entire gravity anomaly to the

presence of dykes, and reduced relief along the density interfaces relative to models without dykes.

The extension across the structures was calculated from the thickness of the feldspathic crust by integrating the fractional crustal thickness anomaly across the structures:

$$\Delta L = \int_{x_1}^{x_2} \left( 1 - \frac{c(x)}{c_0} \right) dx$$

where  $\Delta L$  is the change in length between locations  $x_1$  and  $x_2$ ,  $c(x)$  is the thickness of the feldspathic crust as a function of location, and  $c_0$  is the mean thickness of the crust on either side of the structure. The extension was calculated between the shoulders on either side of the rift for each model, encompassing a zone 131 and 152 km wide for anomalies 1 and 2, respectively. The calculated extension and corresponding extensional strain across the structures for each of the models are given in Extended Data Table 1. The models with an isostatic ratio between the relief at the top and bottom of the feldspathic crust predict greater extension because a larger fraction of the gravity signal is downward continued to the crust–mantle interface, resulting in greater amplification of the short-wavelength anomalies. The extension calculated using the crustal thickness models is an upper bound because some contribution to the gravity anomaly arising from the mechanical or thermal reduction of the crustal porosity beneath the mare load and surrounding the intruded dykes is likely.

We next inverted the Bouguer gravity over the PKT border structures for the best-fit dykes using a Monte Carlo approach. The sources of the anomalies were represented as density anomalies with elliptical cross-sections in the vertical plane perpendicular to the long axes of the anomalies, of assumed density contrast and bottom depth and unknown width and top depth. The bottom depths were set to the typical crustal thickness within the PKT of  $\sim 25 \text{ km}$  (ref. 2), and the density contrasts were set to  $550 \text{ kg m}^{-3}$ , corresponding to a crustal density of  $2,550 \text{ kg m}^{-3}$  (ref. 2) and an intrusion density of  $3,100 \text{ kg m}^{-3}$  (ref. 34). Dykes with elliptical cross-sections were then constructed from a large number of rectangular prismatic elements, and the gravity anomaly was calculated from those prisms<sup>40</sup>. The best-fit solutions were found using a simple Markov chain Monte Carlo (MCMC) approach<sup>14</sup>. The one-standard-deviation ( $1\sigma$ ) confidence intervals on the best-fit solutions were obtained by using a Metropolis–Hastings MCMC to test 20,000 models and analysing the histograms of the resultant model parameters<sup>14</sup>. If the volume of the dyke is accommodated solely by horizontal extension, then the resulting extensions for anomalies 1 and 2 are 21 km and 20 km, respectively, given intrusion into a 25-km-thick crust.

**Thermal modelling.** The thermal evolution of the PKT was modelled following earlier work by Wiczorek and Phillips<sup>10</sup> and Grimm<sup>11</sup>, under the assumption of conductive heat transfer through the mantle. The results of this work are primarily sensitive to the temperatures in the lithosphere, which are dominated by the concentration of heat-producing elements in the crust and the conductive heat transfer through the lithosphere. Although early convection beneath the PKT was possible<sup>12</sup>, this convection would have had only a second-order effect on the temperatures in the lithosphere. We used a finite difference approach to solve the spherical axisymmetric thermal diffusion equation. The model was benchmarked against the analytic solution for half-space cooling from an instantaneous temperature change applied to the surface, as well as by comparison with the results of previous work<sup>10</sup>. The model nodes were divided into crust, mantle, and KREEP components.

The PKT was represented by a spherical cap  $2,000 \text{ km}$  ( $66^\circ$ ) in diameter in which the concentration of heat-producing elements was enhanced. The lack of similarly high concentrations of heat-producing elements on the farside is supported by the lack of evidence for KREEP-rich material within or surrounding the South Pole–Aitken impact basin<sup>3</sup>. The cause for this concentration of incompatible elements on the nearside is not known, but it may be related to a degree-1 Rayleigh–Taylor instability that arose from the gravitational instability of the dense ilmenite-rich cumulates formed in the late stages of magma ocean crystallization<sup>35</sup>. The crustal thickness was set to a uniform value of 40 km in order to isolate the effect of the concentration of heat-producing elements in the PKT. The effect of the thicker crust outside the PKT is less than the uncertainties in the concentration of heat-producing elements and the thermal conductivity of the crust and PKT. We assumed a thermal conductivity of  $2.0 \text{ W m}^{-1} \text{ K}^{-1}$  for the crust and KREEP-rich material, and  $3.0 \text{ W m}^{-1} \text{ K}^{-1}$  for the mantle. The densities of the crust/PKT and mantle were set to  $2,550$  and  $3,200 \text{ kg m}^{-3}$ , respectively, and a specific heat of  $1,200 \text{ J kg}^{-1} \text{ K}^{-1}$  was assumed for all materials.

Previous studies favoured a 10-km-thick layer of KREEP basalt at the base of the crust<sup>10,12</sup>, but other workers have argued that this scenario is not compatible with the gravity and topography of the region and generates too much melt<sup>11</sup>. In our nominal model, we included a 10-km-thick layer of KREEP basalt at the base of the crust. We also considered the case of a 10-km-thick layer of KREEP basalt distributed

uniformly throughout a 40-km-thick crust. We assumed a U concentration in the KREEP basalt of 3.4 p.p.m. by weight, and concentrations of 0.14 p.p.m. and 6.8 p.p.b. in the crust and mantle, respectively<sup>10,12</sup>. We assumed a K/U ratio of 2,500 and a Th/U ratio of 3.7 in all materials<sup>12</sup>. The enhanced concentration of KREEP is given an abrupt edge in the thermal model for simplicity. The thermal effects of this edge are broadened over the thermal diffusion length scale ( $\sim 50$  km for 100 Myr), whereas the stress effects are spread out over a distance comparable to the flexural half-wavelength ( $\sim 540$  km for a lithosphere thickness of 50 km). The overall stress pattern would be unaffected by tapering the margins of the KREEP terrane over length scales of this order. The effects of melting and melt extraction on the temperature evolution were neglected. Extraction of melt would reduce the magnitude of the thermal anomaly in early time steps and decrease the amount of cooling by a modest amount, but would not change the character of the results.

High temperatures throughout the lunar interior are expected after accretion and solidification of the magma ocean<sup>37</sup>. The model was initialized with an approximation to an adiabatic temperature gradient throughout the model domain<sup>10</sup>, increasing linearly from 1,450 K at the surface to 1,500 K at the core–mantle boundary at a radius of 438 km. This temperature profile represents the temperature at the end of an early convective period. In the absence of an early period of convection, the temperatures at the top of the mantle after magma ocean overturn would have been similar<sup>37</sup>. The top boundary condition was set to a constant temperature of 250 K, approximating the radiative equilibrium temperature of the lunar surface. A constant heat flux of 0 was applied as the basal boundary condition at the core–mantle boundary. The model begins at time  $t = 0$  (4.5 Gyr ago) and was run forward in time for 4.5 Gyr. The change in temperature with time was calculated between 4.0 Gyr ago (somewhat before the onset of the geological record) and 3.0 Gyr ago, bracketing the period during which the majority of the maria formed<sup>17,41,42</sup>. It is only the change in temperature that generates thermal stresses in the lithosphere, so even though the PKT was always warmer than its surroundings, its time evolution was characterized by net cooling and thermal contraction because it cooled at a faster rate. The temperature change of the PKT relative to the surroundings was also calculated for illustration purposes by subtracting the temperature change profile at the antipode of the PKT. The absolute change in temperature was used in all stress modelling, but the relative temperature change serves to highlight the evolving thermal anomaly beneath the PKT.

The changes in temperature as functions of time at 25 km depth (the midplane of the 50-km-thick lithosphere assumed for the stress modelling) both within and outside the PKT are shown in Extended Data Fig. 4. Both scenarios for the distribution of KREEP-rich material show similar patterns, but the model with an isolated KREEP-rich layer beneath the crust experiences an early phase of warming in the first few hundred million years. Between 4.0 and 3.0 Gyr ago, both models predict substantially more cooling in the PKT than elsewhere. The mantle immediately below the PKT follows a similar pattern of cooling with time as a result of the decline in heat production within the PKT. In contrast, the mantle at deeper levels warms up as it slowly comes into thermal equilibrium with the overlying KREEP material<sup>10</sup> (Extended Data Fig. 5). However, the net effect of the cooling upper mantle and warming lower mantle approximately cancel out. The temperature changes predicted here are somewhat larger than those of Wiczorek and Phillips<sup>10</sup> as a result of the different ratios between the concentrations of heat-producing elements<sup>12</sup> and the neglect of latent heat and melt extraction effects in this study. Reducing the concentration of radiogenic isotopes or taking into account melt extraction would reduce the magnitudes of the predicted temperature changes and stresses, but would not affect their spatial patterns.

There is substantial uncertainty in the early thermal state of the Moon. The variation of temperature with depth after accretion and solidification of the magma ocean depends strongly on the timescale of accretion<sup>43</sup>, the depth of the magma ocean<sup>21,43</sup>, and the possible gravitational overturn of the magma ocean cumulates<sup>37,38</sup>. However, our models depend primarily on the temperatures within the lithosphere, which are dominated by the time evolution of the heat production within the crust. By 4.0 Gyr ago, the time at which we begin tracking the temperature changes to calculate the strain, the effect of the assumed initial condition on the temperatures in the lithosphere is greatly reduced. The early period of thermal equilibration of the lithosphere is reflected in the  $\sim 200$ -Myr period of increasing temperature for the case of KREEP-rich material concentrated at the base of the crust (Extended Data Fig. 4). The magnitude of this warming is substantially less than the magnitude of the cooling that follows. The possible persistence of mantle convection throughout the time period of interest<sup>12</sup> would affect the distribution of temperature with depth in the mantle but would have little effect on the time evolution of the temperature in the lithosphere.

Both Apollo seismic observations<sup>44</sup> and GRAIL gravity measurements<sup>2</sup> indicate that the Moon's upper crust is fractured and porous, possibly to a depth of  $\sim 20$  km. This porosity is likely to reduce the thermal conductivity of the upper crust<sup>45</sup>. The viscous closure of porosity is a thermally activated process<sup>246</sup>, so the higher temperatures

within the PKT (Extended Data Fig. 4) may have decreased the crustal porosity and increased the thermal conductivity in the PKT relative to its surroundings. This increased thermal conductivity would have acted to accelerate the cooling of the PKT relative to that shown in Extended Data Figs 4 and 5. We have not attempted to model this process in detail, but we note that it will positively reinforce the thermal evolution discussed here.

**Stress modelling.** The stresses resulting from the changes in temperature with time were modelled using the Tekton finite element software<sup>20</sup> in a spherical axisymmetric geometry subject to a uniform radial gravitational acceleration. In order to provide adequate spatial resolution in the PKT, the model domain was limited to the elastic lithosphere, assumed to be 50 km thick (see discussion below). The bottom boundary condition represented the restoring force of the mantle with a pressure that varied with depth, whereas elements were free to move in both vertical and horizontal directions. The effects of the buoyant upward pressure arising from thermal anomalies in the mantle below the PKT were applied to the bottom boundary as an additional pressure term that varied with location on the basis of the thermal model results. This pressure term was calculated as the depth integral of the density contrast relative to background density, scaled by the gravitational acceleration. Although considerable thermal anomalies are predicted in the sub-lithospheric mantle beneath the PKT, the effects of the cooling upper mantle and warming lower mantle largely cancel out. The remaining pressure contributes to a broad upwarping of the surface<sup>11</sup> but has little effect on the short-wavelength stresses that are the focus of this analysis. The final topography and gravity anomalies over the PKT as a whole would have been strongly affected by the flexural resistance of the lithosphere<sup>11</sup>, the thinning of the crust within the PKT<sup>2</sup>, and loading by the maria<sup>12</sup>. The excess basal pressure far from the PKT, representing the effects of net global expansion or contraction, was subtracted from the basal pressure condition throughout the model. The net volume change of the interior could add a uniform compressional or extensional horizontal stress to the lithosphere, depending on the early thermal history of the Moon<sup>14,21,43</sup>. The model domain of a 50-km-thick lithosphere stretching from pole to pole was divided into 600 nodes in the azimuthal direction and 20 nodes in the radial direction, resulting in element dimensions of 9.1 by 2.5 km at the surface.

The predicted temperature change between 4.0 and 3.0 Gyr ago (Extended Data Fig. 5b, d) was used to calculate the resulting instantaneous elastic stresses in the model elements before any deformation<sup>47</sup>:

$$\sigma = \alpha_v \Delta T \frac{E}{3(1-2\nu)}$$

where  $\sigma$  is the stress (taken here to be isotropic),  $\alpha_v$  is the volumetric coefficient of thermal expansion (assumed to be  $2 \times 10^{-5} \text{ K}^{-1}$ ),  $E$  is Young's modulus (assumed to be 100 GPa, which is probably appropriate for the lower crust in which the greatest contraction occurs), and  $\nu$  is Poisson's ratio (assumed to be 0.25). These pre-strain thermal stresses were added to the lithostatic stresses for the initial condition for the finite element model. Imposing the effects of thermal contraction with the pre-strain stresses allows the resultant deformation and its effects on the stress field to arise self-consistently in the model.

The elastic stresses were calculated relative to the far-field values at the opposite side of the Moon in order to isolate the effects of thermal contraction of the PKT. Geological and geophysical evidence suggests that the net stress state of the Moon may have evolved from global expansion and extension to contraction and compression over the course of its thermal evolution<sup>14,21</sup>. In this scenario, the net global stress change at the time of formation of the border anomalies may have been small. However, theoretical models have shown that an early period of global expansion is difficult to generate for many likely lunar formation scenarios<sup>43</sup>. We put this question aside and focused instead on the local stresses within and surrounding the PKT relative to the typical stresses far from the region. These stresses would have been modified by the global stress state at the time of interest by the addition of a uniform compressional or extensional horizontal stress. In addition to the relative stresses, we also show the difference between the in-plane horizontal (that is, radial to the centre of the PKT) and vertical stresses ( $\sigma_h - \sigma_v$ ) and the deviatoric horizontal stress ( $\sigma_h - \sigma_p$ ), where  $\sigma_p$  is the pressure or mean stress value over all three directions (Extended Data Fig. 6). The width of the zone of predicted extension ( $\sim 400$  km) is somewhat wider than the observed border structures ( $\sim 200$  km), but localization of the strain would probably have occurred if the structures are analogous to lava-flooded rifts. Similar stress patterns are predicted if KREEP-rich material is distributed uniformly through the crust, though the magnitudes of the stresses are reduced (Extended Data Fig. 7) because of the reduced temperature changes (Extended Data Fig. 5c, d).

The stresses predicted by the model are dominated by the simple horizontal contractional stresses within the lithosphere. However, volumetric contraction also induces small changes to the surface topography, which generates bending stresses of small magnitude. Models in which the vertical displacement was set to zero at either

the top or the bottom of the model domain resulted in similar stress fields, demonstrating that bending stresses do not contribute markedly.

The modelling in this study was intentionally simple in order to isolate the effect of the contracting cap within the PKT. This analysis did not consider the effects of spatial or temporal variations in the lithosphere thickness. Because the dominant source of stress is the horizontal contraction of the lithosphere within the PKT, the stresses for the case of a variable lithosphere should be similar. This model represented only the elastic stresses within the lithosphere. A viscoelastic model of the lithosphere and underlying mantle would predict a viscous transition zone at the base of the lithosphere within which the stresses decreased to zero at depth. Coupling of the thermal and viscoelastic evolution would result in a lithosphere that thickens with time, and would probably reduce the magnitude of the predicted extension, but would not change the character of the results. Within the PKT, the stresses are characterized by compression in the upper lithosphere and extension in the lower lithosphere, whereas at the edges of the PKT the extensional stress reaches the surface. However, the frictional strength at the surface should approach 0 MPa, allowing release of the shallow compressional stresses. Brittle compressional failure of the frictionally weak upper lithosphere throughout the PKT would allow further contraction of the spherical cap, substantially enhancing the extension at its margins.

In order to model directly the formation of the observed border structures, it would be necessary to localize the extension through tectonic failure. The localization of the extensional failure at discrete rift zones in the border structures would be dependent on the strain rate, rheology, and crustal thickness<sup>48</sup>. Failure at the edges of the PKT would relieve the stresses in the interior and allow the spherical cap to pull away from the surrounding lithosphere. Future work is needed to model more directly the formation of these border structures. In this work, we simply show that thermal contraction of the PKT predicts extension at its edges, providing a straightforward mechanism for generating the PKT border structures. Additional stresses arising from uplift or subsidence of the lithosphere<sup>11,12</sup> and magmatic processes would have also probably played a role.

Zones favourable to the ascent of magma-filled dykes through the lithosphere were identified as those experiencing in-plane horizontal extension relative to the vertical stress and a favourable vertical stress gradient. Horizontal extension is required for the formation of vertical dykes, which would otherwise flatten out to produce horizontal sills. In addition, the upward propagation of the dykes requires that the vertical gradient in the confining horizontal stress in the lithosphere ( $d\sigma_h/dz$ , where positive stresses are tensile,  $z$  is positive upward, and  $\sigma_h$  includes both the lithostatic stress and the added tectonic stress) be greater than the hydrostatic pressure gradient in the magma, causing the lower tip of the dyke to pinch shut as the upper tip propagates upward. The low density of the lunar crust<sup>2</sup> is an impediment to the rise of magma, even in a neutral stress state. Magma ascent is favoured in cases in which the upper lithosphere is in a state of extension relative to the lower lithosphere<sup>49</sup>. For a magma density of  $2,900 \text{ kg m}^{-3}$ , this state corresponds to a vertical gradient in the horizontal stress in the lithosphere in excess of  $4.7 \text{ MPa km}^{-1}$ . However, the stress gradient in the upper portions of a conductively cooling lithosphere with internal heat production and basal heating is generally not conducive to magma ascent as a result of the increasing horizontal extension with depth caused by the declining thermal gradient in the lithosphere with time. This problem could be ameliorated by a failure-induced reduction in the extensional stresses in the lower crust, by volatile exsolution within the magma to enhance the driving force for magma ascent<sup>50</sup>, or by a pressurized magma reservoir at depth. We use the criterion of a vertical stress gradient  $>4.7 \text{ MPa km}^{-1}$  for unassisted magma rise, and we also look at the stress gradient relative to the far-field value antipodal to the PKT to assess the relative tendency for magma to rise through the lithosphere if assisted by other factors as discussed above.

By these criteria, the zone at the margin of the PKT experiences stresses most conducive to magma ascent and eruption. Extensional horizontal stresses radial to the centre of the PKT would facilitate the formation of circumferential dykes throughout the full vertical extent of the lithosphere. The stress gradient in this zone is more conducive to magma ascent than anywhere else on the Moon. For the case of heating by a layer of KREEP at the base of the crust, magma would be predicted to rise unassisted to the middle of the lithosphere, whereas further ascent would require additional driving forces (Extended Data Fig. 6c). For the case of heating by KREEP distributed throughout the crust, the zone at the edge of the PKT is still the preferred location of magma ascent, but some added driving force such as volatile exsolution or a pressurized magma chamber is required for the rise of magma into the crust (Extended Data Fig. 7c).

The model also predicts changes to the surface topography. The thermal contraction of the lithosphere with time causes surface subsidence due to the vertical component of that contraction. Additionally, the horizontal shortening of the spherical cap centred on the PKT results in further subsidence because the decrease in the area of the cap must be accommodated by an increase in the radius of curvature,

resulting in a decrease in elevation. Taking into account both the thermal contraction of the lithospheric cap in the PKT and the effects of thermal anomalies in the mantle, our models predict changes in surface topography less than 0.5 km during the period between 4.0 and 3.0 Gyr ago. This result cannot be directly compared with the observed topography because it represents only the change in topography over a fraction of lunar history. However, we note that the predicted elevation changes are smaller than the observed relief. The topographic depression within Procellarum cannot be explained by thermal subsidence alone.

Previous work indicated that the patterns of uplift predicted by thermal models of the PKT are difficult to reconcile with the observed long-wavelength gravity and topography<sup>11,12</sup>. However, the gravity and topography within the PKT are also probably affected by variations in the thickness of the crust and by possible density anomalies in the underlying mantle<sup>12</sup>. The low topography within the PKT may also be affected by a reduction in the porosity of the crust from thermal annealing<sup>46</sup>. For a lunar crustal porosity of 12% (ref. 2), annealing the pore space in the lower 10–20 km would reduce the surface elevation by 1.2–2.4 km, consistent with the observed relief.

**Geometry of the PKT border structures.** In order to understand the shape of the observed pattern of border structures in a spherical geometry, we used the spherical law of cosines to determine the vertex angle  $\theta$  for a regular polygon on the surface of a sphere, with  $n$  sides of angular length  $s$  (in radians):

$$\theta = 2 \arccos \left( \frac{\cos(h) - \cos(h) \cos(s)}{\sin(h) \sin(s)} \right)$$

where  $h$  is the angular length of the path from the polygon centre to the vertex, given by:

$$h = \arcsin \left( \sqrt{\frac{\cos(s) - 1}{\cos(2\pi/n) - 1}} \right)$$

These equations were used to calculate the side length at which a regular polygon with  $120^\circ$  vertices will have either four or five sides, rather than the six-sided figure expected for a flat Euclidean geometry.

We suggest that the quasi-rectangular pattern of border structures surrounding the PKT is consistent with the intersection of linear rifts at  $120^\circ$ -angle triple junctions when the effect of the curvature of the surface is taken into account. Although the PKT border structures display some intermediate kinks and intersections, the overall pattern is quasi-rectangular. Similarly, although small-scale contraction-crack polygons of all types in nature often have highly irregular forms, in the absence of competing effects (such as progressive subdivision of the polygons) the average structure is hexagonal because of the dominance of  $120^\circ$ -angle triple junctions at the vertices<sup>51</sup>. At small scales, the diameter of the polygons is determined by the size of the stress shadow around the fractures, which is proportional to the depth to which the fractures propagate<sup>51</sup>. The depth of fracturing for small contraction-crack polygons on Earth is dictated by the strain rate, the variation of stress with depth, and the rheology of the material in which the fractures form<sup>51</sup>. For the PKT, the size of the polygon was probably determined instead by the diameter of the tensile stress belt at the surface surrounding the thermal anomaly. The propagation of the fractures or rifts into the interior of the region may have been prevented by the compressional stresses in the upper lithosphere above the thermal anomaly (Extended Data Figs 6, 7), which may have had an effect similar to the stress shadows around fractures in small-scale polygons. Although the analogy of the quasi-rectangular pattern of border structures to smaller polygonal fracture patterns provides a simple explanation for the observed geometry, further testing of this hypothesis will require consideration of the competing effects of the regional stress directions, the stress field generated by the structures themselves, and the concentration of stress at the tips of the propagating faults or dykes. An alternative explanation for the pattern of border structures that warrants further consideration is that the distribution of KREEP-rich material in the subsurface may follow a quasi-rectangular pattern. However, the distribution of KREEP-rich material in the subsurface is poorly constrained, and the distribution on the surface is strongly affected by the ejecta of the Imbrium basin and the distribution of KREEP-rich maria (which was controlled in part by the pattern of the PKT border structures).

**Parallels between the PKT and the south polar terrain of Enceladus.** The overall pattern of the PKT border structures bears a strong resemblance to that of the border structures surrounding the south polar terrain (SPT) of Saturn's icy moon Enceladus, which are also quasi-rectangular in outline<sup>25</sup>. However, as discussed in the main text and expanded upon below, substantial differences exist between these provinces and their inferred evolutionary histories. We emphasize that we do not suggest that the specific processes and evolutionary paths of these two regions were identical. Rather, the gross similarities between these two provinces on different bodies suggests broad parallels in the processes governing their evolution. Here we



summarize the basic properties of each province, and then discuss the SPT in more detail.

The PKT on the Moon is a broad area of enhanced surface heat flow as a result of the high concentrations of heat-producing elements within the KREEP-rich material<sup>3,10</sup> (Extended Data Fig. 8). This compositional anomaly is probably a result of the concentration beneath the nearside of the late-stage crystallization products of the lunar magma ocean<sup>35</sup>, including dense ilmenite-rich cumulates and KREEP-rich material with high concentrations of U, Th, and K. The PKT was the most volcanically active region on the Moon and contains the majority of the mare basalt provinces<sup>10</sup>. GRAIL gravity anomalies and gradients indicate that the PKT is surrounded by a quasi-rectangular set of magmatic-tectonic structures with straight sides and angular intersections. The border anomalies along the northern (Mare Frigoris) and eastern edges of the PKT occur beneath maria that are confined within elongated topographic depressions, whereas the border anomalies on the western and southern edges of the PKT lie adjacent and interior to the topographic step up to the highlands. The PKT is characterized by low topography that is largely isostatically compensated at long wavelengths. This compensated depression can be explained by a crust that is thinner<sup>2</sup> or denser than that of surrounding areas, by the presence of denser materials at depth, or by a combination of these effects. Thermal annealing of the pore space beneath the PKT due to its high heat flow may have increased the bulk density of the crust at depth, which may contribute to the low topography<sup>46</sup>. Deeper density anomalies could result from either the intrusion of KREEP-rich magma into the lower crust, or from the presence of a remnant of the ilmenite-rich cumulates in the upper mantle that may not have fully mixed into the deeper interior<sup>38</sup>. Although a thinner crust probably explains most of the observed topography, contributions from reduced crustal porosity and the presence of dense materials within or below the crust appear likely.

The SPT on Enceladus is an area of strongly enhanced surface heat flow<sup>26,52,53</sup> (Extended Data Fig. 8) as a result of either localized tidal heating or the localized release of global tidal heating. The source of this thermal anomaly is thought to be related to the presence of a regional liquid water sea or the regional thickening of a global ocean beneath the SPT, which may be a result of locally enhanced tidal heating and would itself contribute to the enhanced tidal heating<sup>27,54,55</sup>. The SPT is cryovolcanically active, as revealed by the plume of water vapour and icy particles emanating from the parallel 'tiger stripes' fractures in the centre of the SPT<sup>24</sup>. Cassini Imaging Science Subsystem (ISS) images reveal the SPT to be bounded by a quasi-rectangular set of tectonic structures with straight sides and angular intersections<sup>24</sup>. These border structures occur near the edges of the topographic depression containing the SPT<sup>24</sup>, located either at or just above the topographic step leading from the SPT up to the surrounding surface<sup>25</sup> (Fig. 3c). The SPT is characterized by low topography<sup>25,56</sup> that is largely isostatically compensated at long wavelengths<sup>57</sup>, which is best explained by the presence of a relatively dense subsurface ocean<sup>57,58</sup>. Depressions in other areas of Enceladus<sup>25</sup> have been explained as resulting from the thermal annealing of the pore space due to the presence of local thermal anomalies beneath these regions in the past<sup>59</sup>. Some contribution to the SPT depression from a reduction of the pore space seems probable given the high observed heat flow. However, the large apparent depth of compensation of the SPT indicated by the long-wavelength gravity and topography suggests that the effect of a deeper ocean dominates<sup>57</sup>.

Although there are notable large-scale morphological and geodynamic similarities between the PKT and the SPT, there are many differences between these provinces as well. The thermal anomaly in the SPT is a result of tidal, rather than radiogenic, heating. Multiple mechanisms have been proposed to explain the high heat flux in the SPT, including viscous heating in the ice shell<sup>35</sup> and shear heating along fractures<sup>60</sup>. Each of these mechanisms ultimately relies on tidal energy from the gravitational interaction of Enceladus with Saturn. However, the expected steady-state rate of tidal heating for the present-day eccentricity<sup>61</sup> is not sufficient to maintain the observed heat flux within the SPT<sup>26,53</sup> or the inferred subsurface ocean beneath the region<sup>27</sup>. Recent results have revised the lower bounds on the heat flow downward<sup>52</sup>, but values remain above the expected steady-state tidal heating unless the dissipation within Saturn is higher than expected from theoretical considerations<sup>62</sup>. This discrepancy may be explained if the SPT today is in a transient state of high heat flow following an earlier period of high orbital eccentricity during which the ocean formed<sup>27</sup>. This scenario implies a time-variable heat flux in which the SPT may be cooling today.

The lack of craters within the SPT<sup>24</sup> suggests an earlier episode of volcanic resurfacing, lithosphere recycling<sup>28,63</sup>, or viscous relaxation of the craters<sup>64</sup>. Each of these scenarios could have resulted in a regional thermal anomaly, followed by a period of cooling and contraction of the ice throughout the SPT. Globally, substantial lateral and temporal variations in the heat flux have been inferred on the basis of high local heat fluxes indicated by the relaxation of craters<sup>64</sup> and the flexural support of topography<sup>65</sup>. Structures similar in scale and morphology to the SPT on the leading and trailing hemispheres suggest similar activity at those locations in the past<sup>66</sup>,

further supporting spatial and temporal variability in the thermal state of Enceladus's ice shell.

The SPT border structures are each composed of a belt of closely spaced parallel ridges, surrounded by an inward (southward) facing scarp<sup>24,67</sup>. The ridge belts probably formed by compression<sup>24,25</sup>, though extensional deformation<sup>68</sup> or more complicated scenarios<sup>69</sup> may have played a role in the formation of the south-facing scarps. For the compressional interpretation, it has been proposed that the tectonics in the SPT was driven by regional thermal expansion<sup>70</sup>, which is similar in nature but opposite in sign to what is proposed here for the PKT. At some intersections, the border scarps are continuous with fracture belts extending northward from 120°-angle triple junctions<sup>67</sup>, consistent with an extensional origin for the outer scarp. However, the folded terrains confined within the angular corners are indicative of compressional deformation<sup>24</sup>. Compressional folding is also observed in the interior of the SPT away from the border structures<sup>71</sup>, whereas tensile opening of the 'tiger stripe' fractures is required to explain the observed volcanic venting<sup>72</sup>. Thus, both compressional and extensional tectonics have been active along the border structures and within the interior of the SPT.

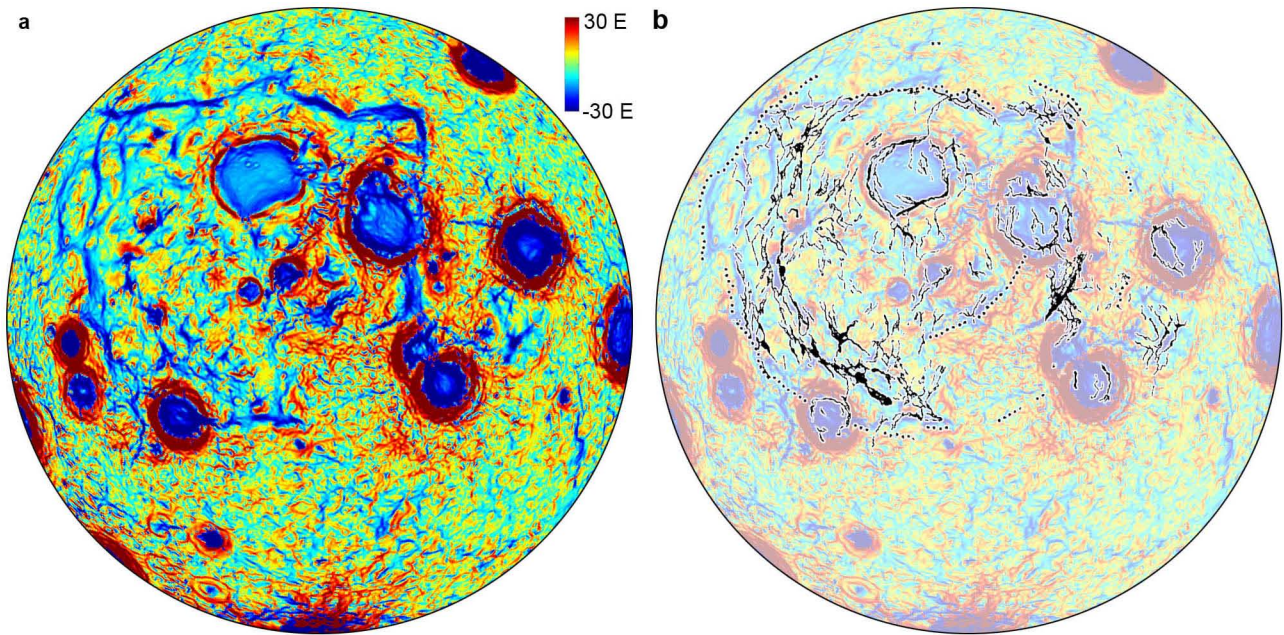
Some structures observed within the SPT are broadly consistent with our model predictions for the PKT. The models predict the upper lithosphere within a cooling lithospheric cap to be in a state of compression due to its coupling with the contracting lower lithosphere, whereas the cap would be surrounded by a belt in which extensional stresses pervade the entire lithosphere (Fig. 4, Extended Data Figs 6, 7). This stress pattern predicts broad compressional deformation of the upper lithosphere within the SPT and lithosphere-scale normal faulting at the margins. However, we emphasize that simple thermal expansion and contraction alone cannot explain the extensive tectonics within the SPT. The extensive tectonic modification and resulting large strains may indicate an earlier period of mobile-lithosphere tectonics<sup>28,63</sup>.

Enceladus is much smaller than the Moon (radii of 252 and 1,738 km, respectively). Although the SPT is much smaller than the PKT in physical size (~300 km versus ~2,000 km), they are similar in angular size (Fig. 3). Thus, the geometric arguments for the formation of the quasi-rectangular PKT border structures due to the intersection of tectonic structures at 120° angles on a spherical surface may have relevance to the SPT as well. The different values of gravitational acceleration at the surfaces of the Moon and Enceladus would not directly affect their thermal evolution, but would have had an impact on the ensuing tectonic and volcanic processes.

Thus, we suggest that broadly similar geodynamic processes may have been at work in the PKT and the SPT. Both regions are characterized by strong thermal anomalies, enhanced volcanic activity, and low topography. The quasi-rectangular structures surrounding both provinces are consistent with the expected shapes of sets of tectonic structures intersecting at 120°-angle triple junctions. However, the specific evolutionary paths of the provinces were probably substantially different as a result of the differences in the sources of heat, temporal variations in heat flux, and rheologies of the lithospheres. Our current understanding of the formation and evolution of both structures is incomplete. Nevertheless, the two provinces highlight the important effect that regional thermal anomalies can have on the volcanic and tectonic evolution of quite different planetary bodies.

1. Lemoine, F. G. *et al.* GRGM900C: A degree 900 lunar gravity model from GRAIL primary and extended mission data. *Geophys. Res. Lett.* **41**, 3382–3389 (2014).
2. Reed, G. B. *Application of Kinematical Geodesy for Determining Short Wave Length Components of the Gravity Field by Satellite Gradiometry* (Tech. Rep. 201, Department of Geodetic Science, The Ohio State University, 1973).
3. Wieczorek, M. A. & Phillips, R. J. Potential anomalies on a sphere: applications to the thickness of the lunar crust. *J. Geophys. Res.* **103**, 1715–1724 (1998).
4. Kiefer, W. S. *et al.* The density and porosity of lunar rocks. *Geophys. Res. Lett.* **39**, L07201 (2012).
5. Parmentier, E. M., Zhong, S. & Zuber, M. T. Gravitational differentiation due to initial chemical stratification: origin of lunar asymmetry by the creep of dense KREEP? *Earth Planet. Sci. Lett.* **201**, 473–480 (2002).
6. Hess, P. C. & Parmentier, E. M. A model for the thermal and chemical evolution of the Moon's interior: implications for the onset of mare volcanism. *Earth Planet. Sci. Lett.* **134**, 501–514 (1995).
7. Elkins-Tanton, L. T., Burgess, S. & Yin, Q.-Z. The lunar magma ocean: reconciling the solidification process with lunar petrology and geochronology. *Earth Planet. Sci. Lett.* **304**, 326–336 (2011).
8. Elkins-Tanton, L. T., Van Orman, J. A., Hager, B. H. & Grove, T. L. Re-examination of the lunar magma ocean cumulate overturn hypothesis: melting or mixing is required. *Earth Planet. Sci. Lett.* **196**, 239–249 (2002).
9. DeHon, R. A. Maximum thickness of materials in the western mare basins. *Proc. Lunar Planet. Sci. Conf.* **9**, 229–231 (1978).
10. Blakely, R. J. *Potential Theory in Gravity and Magnetic Applications* (Cambridge Univ. Press, 1995).
11. Hiesinger, H., Jaumann, R., Neukum, G. & Head, J. W. Ages of mare basalts on the lunar nearside. *J. Geophys. Res.* **105**, 29239–29275 (2000).

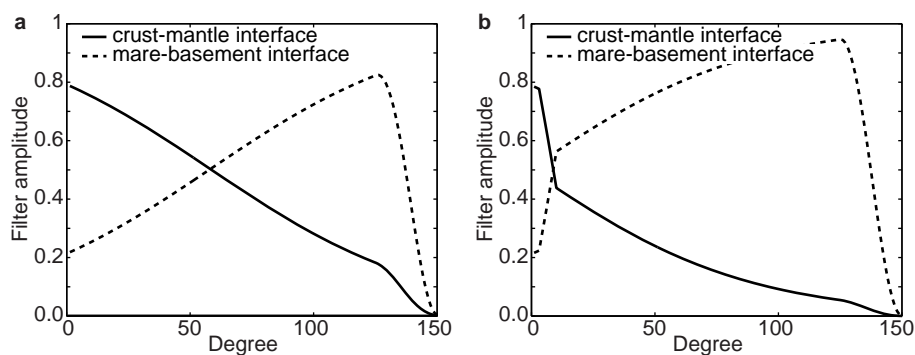
42. Hiesinger, H. *et al.* Ages and stratigraphy of mare basalts in Oceanus Procellarum, Mare Nubium, Mare Cognitum, and Mare Insularum. *J. Geophys. Res.* **108**, 5065 (2003).
43. Pritchard, M. E. & Stevenson, D. J. in *Origin of the Earth and Moon* (eds Canup, R. M. & Righter, K.) 179–196 (Univ. Arizona Press, 2000).
44. Toksöz, M. N., Dainty, A. M., Solomon, S. C. & Anderson, K. R. Structure of the Moon. *Rev. Geophys. Space Phys.* **12**, 539–567 (1974).
45. Presley, M. A. & Christensen, P. R. Thermal conductivity measurements of particulate materials: 4. Effect of bulk density for granular particles. *J. Geophys. Res.* **115**, E07003 (2010).
46. Besserer, J. *et al.* Theoretical and observational constraints on lunar mega-regolith thickness. *Lunar Planet. Sci.* **44**, abstr. 2463 (2013).
47. Turcotte, D. L. & Schubert, G. *Geodynamics* 2nd edn (Cambridge Univ. Press, 2001).
48. Buck, W. R. Modes of continental lithospheric extension. *J. Geophys. Res.* **96**, 20161–20178 (1991).
49. McGovern, P. J. *et al.* Impact-generated loading and lithospheric stress gradients at lunar impact basins: implications for maria emplacement scenarios. *Lunar Planet. Sci.* **44**, abstr. 3055 (2013).
50. Wilson, L. & Head, J. W. Ascent and eruption of basaltic magma on the Earth and Moon. *J. Geophys. Res.* **86**, 2971–3001 (1981).
51. Lachenbruch, A. H. Mechanics of thermal contraction cracks and ice-wedge polygons in permafrost. *Spec. Pap. Geol. Soc. Am.* **70**, 1–66 (1962).
52. Spencer, J. R. *et al.* Enceladus heat flow from high spatial resolution thermal emission observations. *EPSC Abstr.* **8**, EPSC2013–2840–2011 (2013).
53. Spencer, J. R. *et al.* Cassini encounters Enceladus: background and the discovery of a south polar hot spot. *Science* **311**, 1401–1405 (2006).
54. Roberts, J. H. & Nimmo, F. Near-surface heating on Enceladus and the south polar thermal anomaly. *Geophys. Res. Lett.* **35**, L09201 (2008).
55. Tobie, G., Cadek, O. & Sotin, C. Solid tidal friction above a liquid water reservoir as the origin of the south pole hotspot on Enceladus. *Icarus* **196**, 642–652 (2008).
56. Thomas, P. C. *et al.* Shapes of the Saturnian icy satellites and their significance. *Icarus* **190**, 573–584 (2007).
57. Iess, L. *et al.* The gravity field and interior structure of Enceladus. *Science* **344**, 78–80 (2014).
58. Collins, G. C. & Goodman, J. C. Enceladus' south polar sea. *Icarus* **189**, 72–82 (2007).
59. Besserer, J., Nimmo, F., Roberts, J. H. & Pappalardo, R. T. Convection-driven compaction as a possible origin of Enceladus's long wavelength topography. *J. Geophys. Res. Planets* **118**, 908–915 (2013).
60. Nimmo, F., Spencer, J. R., Pappalardo, R. T. & Mullen, M. E. Shear heating as the origin of the plumes and heat flux on Enceladus. *Nature* **447**, 289–291 (2007).
61. Meyer, J. & Wisdom, J. Tidal heating in Enceladus. *Icarus* **188**, 535–539 (2007).
62. Lainey, V. *et al.* Strong tidal dissipation in Saturn and constraints on Enceladus' thermal state from astrometry. *Astrophys. J.* **752**, 14 (2012).
63. Barr, A. C. Mobile lid convection beneath Enceladus' south polar terrain. *J. Geophys. Res.* **113**, E07009 (2008).
64. Bland, M. T., Singer, K. N., McKinnon, W. B. & Schenk, P. M. Enceladus' extreme heat flux as revealed by its relaxed craters. *Geophys. Res. Lett.* **39**, L17204 (2012).
65. Giese, B. *et al.* Enceladus: an estimate of heat flux and lithospheric thickness from flexurally supported topography. *Geophys. Res. Lett.* **35**, L24204 (2008).
66. Crow-Willard, E. & Pappalardo, R. T. Global geological mapping of Enceladus. *EPSC Abstr.* **6**, EPSC–DPS2011–2635–2011 (2011).
67. Spencer, J. R. *et al.* in *Saturn from Cassini-Huygens* (eds Dougherty, M. K., Esposito, L. W. & Krimigis, S. M.) 683–724 (Springer, 2009).
68. Walker, C. C., Bassis, J. N. & Liemohn, M. W. On the application of simple rift basin models to the south polar region of Enceladus. *J. Geophys. Res.* **117**, E07003 (2012).
69. Yin, A. & Pappalardo, R. T. Left-slip faulting along the tiger stripe fractures: implications for the tectonic evolution of the south polar terrain, Enceladus. *Lunar Planet. Sci.* **44**, abstr. 1145 (2013).
70. Gioia, G., Chakraborty, P., Marshak, S. & Kieffer, S. Unified model of tectonics and heat transport in a frigid Enceladus. *Proc. Natl Acad. Sci. USA* **104**, 13578–13581 (2007).
71. Barr, A. C. & Preuss, L. J. On the origin of the south polar folds on Enceladus. *Icarus* **208**, 499–503 (2010).
72. Hurford, T. A. *et al.* Eruptions arising from tidally controlled periodic openings of rifts on Enceladus. *Nature* **447**, 292–294 (2007).



**Extended Data Figure 1 | Comparison of the GRAIL gravity gradients with proposed Procellarum basin ring structures.** **a**, Bouguer gravity gradients (in units of Eötvös;  $1 \text{ E} = 10^{-9} \text{ s}^{-2}$ ) on a Lambert azimuthal equal-area

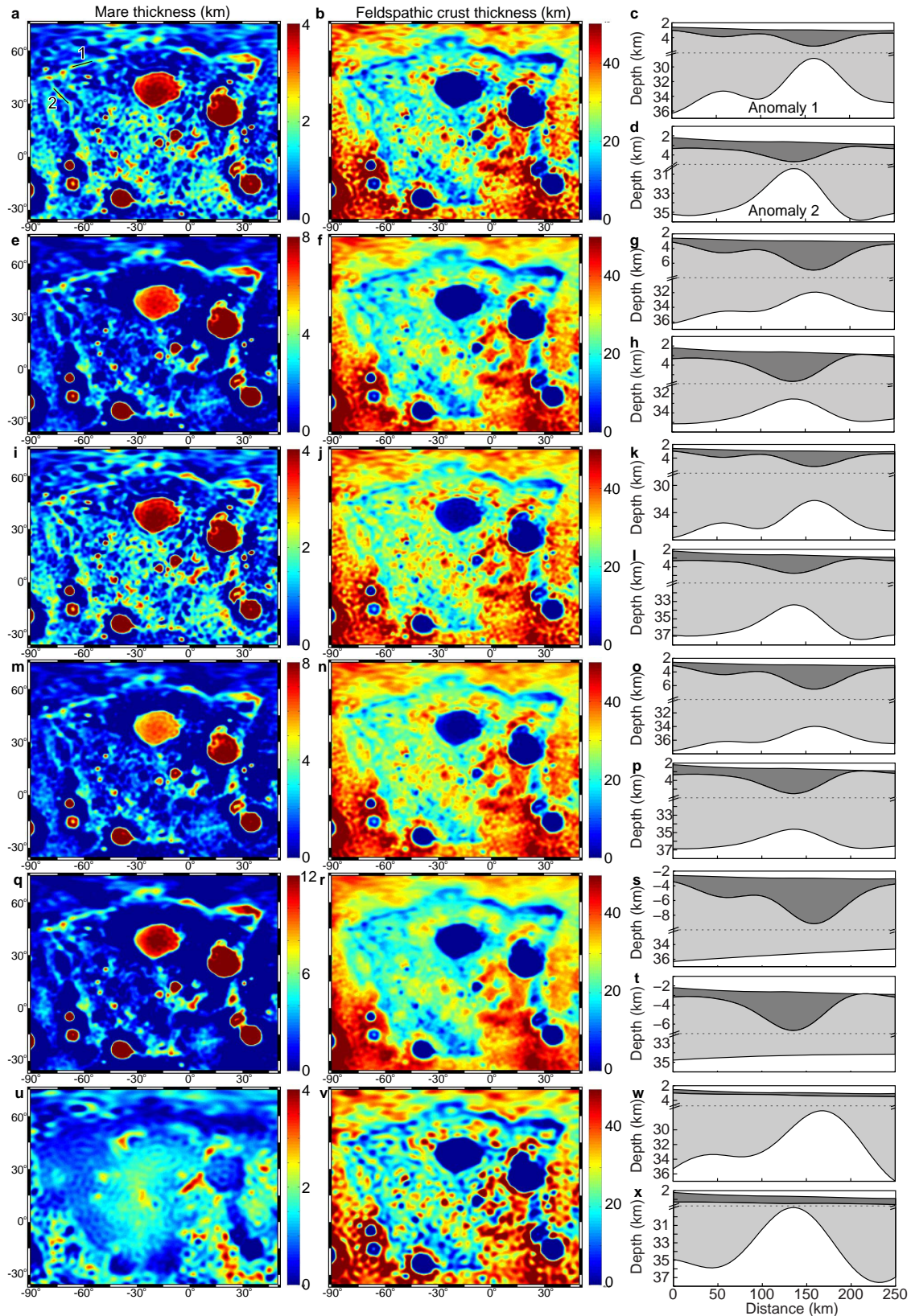
projection of the nearside of the Moon. **b**, Muted gravity gradients overlaid with mapped mare boundaries and scarps (dots) and wrinkle ridges (lines). Modified from figure 1 of ref. 5 with permission.





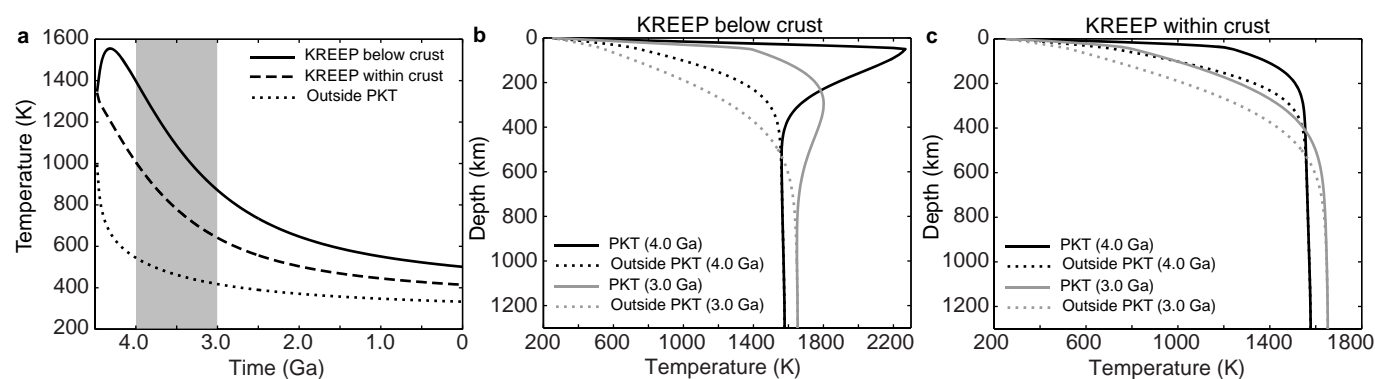
**Extended Data Figure 2 | Amplitudes of filters applied during the crustal thickness modelling.** **a, b,** Filters were applied during the calculation of the relief along the crust–mantle interface (solid lines) and the mare–basement interface (dashed lines) for cases in which the relief along the two interfaces was either isostatic before mare loading (**a**) or equal and opposite in amplitude (**b**). The filter in **b** imposes the isostatic condition from degrees 1 to 3 and a linear

transition to the equal-amplitude filter from degrees 3 to 10. Both filters apply a cosine taper from degrees 125 to 150. The mare–basement filter is shown for illustration purposes only. In practice, the relief along the mare–basement interface was calculated from the residual Bouguer anomaly after the calculation of the crust–mantle interface relief (equivalent to using the filter shown with the original Bouguer gravity).



**Extended Data Figure 3 | Predicted thicknesses of the crust and maria and average cross-sections across two of the border anomalies.** Predicted thickness of the maria (left column) and underlying feldspathic crust (middle column), and cross-sections of the modelled structures of anomaly 1 (right column, top) and anomaly 2 (right column, bottom) showing the variations in the thicknesses of the mare (dark grey) and feldspathic crust (light grey). Models are for cases as follows: **a–d**, isostatic relief along the two interfaces

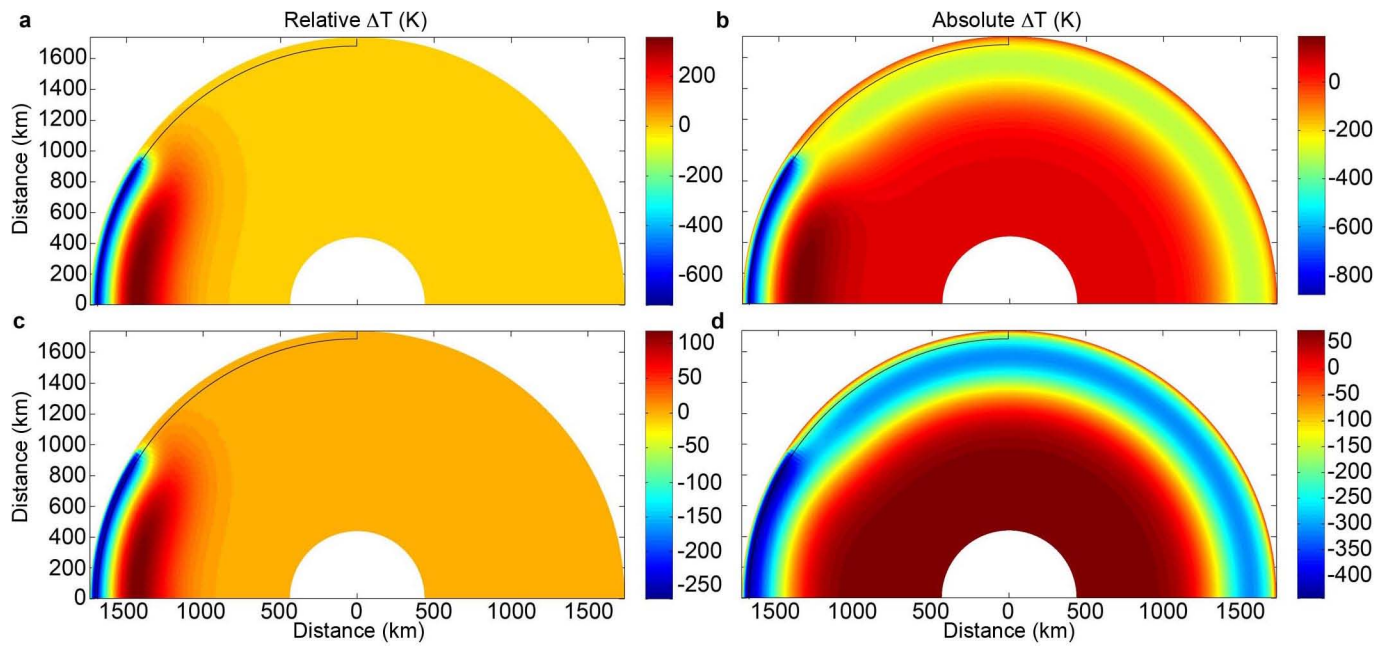
before mare infilling with a mantle density of  $3,220 \text{ kg m}^{-3}$ ; **e–h**, equal-amplitude relief along the two interfaces with a mantle density of  $3,220 \text{ kg m}^{-3}$ ; **i–l**, isostatic relief along the two interfaces before mare infilling with a mantle density of  $3,500 \text{ kg m}^{-3}$ ; **m–p**, equal-amplitude relief along the two interfaces with a mantle density of  $3,500 \text{ kg m}^{-3}$ ; **q–t**, all gravity anomalies at degrees  $>10$  ascribed to relief on the mare–basement interface; and **u–x**, all gravity anomalies at degrees  $>10$  ascribed to relief on the crust–mantle interface.



**Extended Data Figure 4 | Temperature evolution within and outside the PKT.** **a**, The temperatures as functions of time at a depth of 25 km are shown within the PKT for cases in which KREEP-rich material is either concentrated at the base of the crust (solid line) or is distributed throughout the crust (dashed line), as well as the temperature outside the PKT (dotted line).

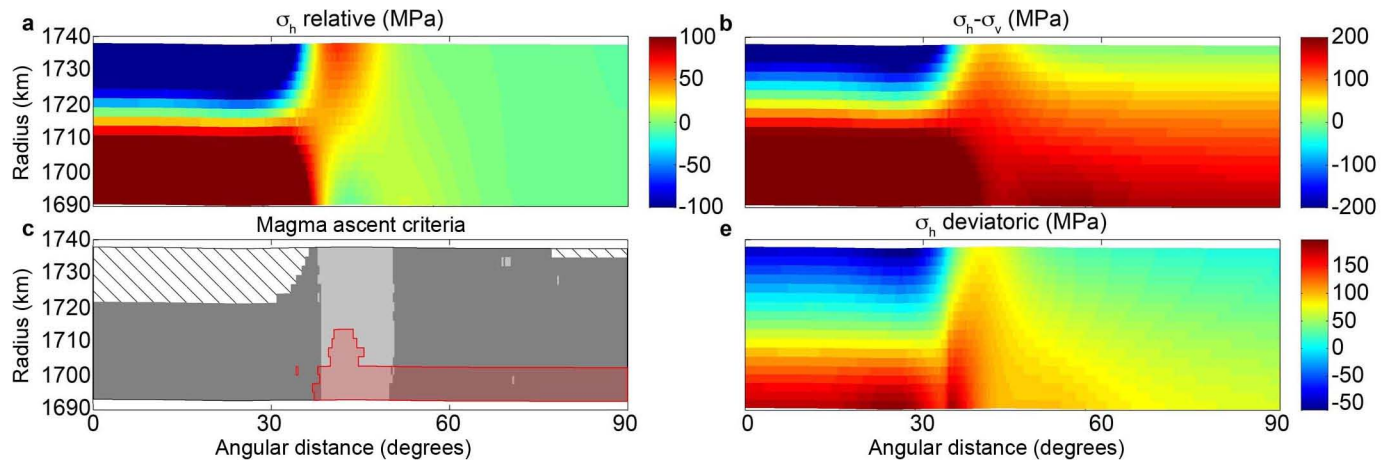
The period between 4.0 and 3.0 Gyr ago that is the focus of the stress modelling is indicated by the shaded box. **b**, **c**, The temperatures as functions of depth both inside and outside the PKT are shown for KREEP-rich material concentrated at the base of the crust (**b**) and for KREEP-rich material distributed throughout the crust (**c**).





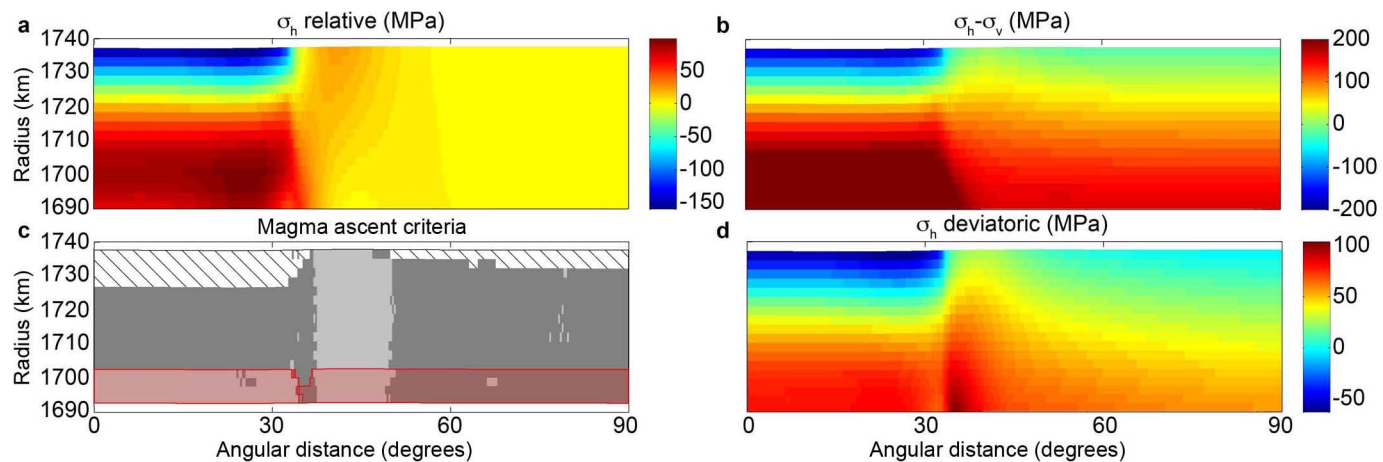
**Extended Data Figure 5 | Predicted changes in temperature relative to areas outside the PKT and absolute temperature change between 4.0 and 3.0 Gyr ago.** Results are shown for cases with KREEP concentrated at the base of the crust (a, b) and KREEP distributed throughout the crust (c, d). The PKT is

centred on the pole at the left side of the panels. The region shown in Extended Data Figs 6 and 7 (encompassing  $90^\circ$  of arc extending radially outward from the centre of the PKT and downward to a depth of 50 km) is outlined in black.



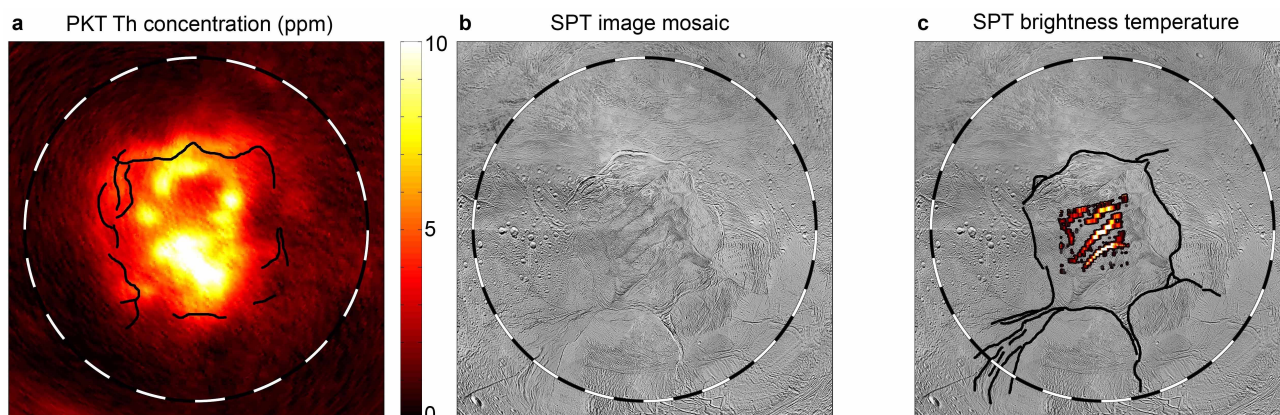
**Extended Data Figure 6 | Predicted lithospheric stresses and magma ascent for the case of 10 km of KREEP at the base of the crust.** Cross-sections show the following: **a**, the in-plane horizontal stresses (radial to the centre of the PKT, the far-field stress profile was subtracted to calculate the relative stress); **b**, the difference between the in-plane horizontal stress and the vertical stress; **c**, the magma ascent criteria; and **d**, the deviatoric stress. The magma ascent

criteria in **c** reveal portions of the crust in which the horizontal stresses are tensile relative to the vertical stresses to permit the formation of vertical dykes (dark grey), where the vertical stress gradient is more favourable to magma ascent than the lithosphere far from the PKT (light grey), where magma will rise unassisted by other factors such as pressurized magma chambers (red), and where none of the criteria are satisfied (diagonal lines).



**Extended Data Figure 7 | Predicted lithospheric stresses and magma ascent for the case of 10 km of KREEP basalt distributed uniformly through a 40-km-thick crust.** All panels are as for Extended Data Fig. 6.





**Extended Data Figure 8 | Additional comparisons of Procellarum KREEP terrane to the Enceladus south polar terrain (SPT).** **a**, The PKT is characterized by high heat flow as a result of the enhanced abundances of radioactive elements<sup>3</sup> (represented by the concentration of thorium<sup>4</sup>). **b**, The border structures of the SPT as revealed by Cassini ISS images<sup>24</sup> also trace a

quasi-rectangular pattern enclosing a region of elevated brightness temperatures and enhanced heat flow<sup>26</sup> (**c**) All maps are in a simple polar projection. In all panels, the circle corresponds to an angular diameter of 180° of surface arc, divided into 10° increments.

Extended Data Table 1 | Extension and strain across two border anomalies

Filter	$\rho_m$ (kg/m <sup>3</sup> )	Anomaly 1		Anomaly 2	
		extension	strain	extension	strain
isostatic	3220	15 km	0.11	13 km	0.08
equal amplitude	3220	12 km	0.09	10 km	0.07
isostatic	3500	11 km	0.09	10 km	0.06
equal amplitude	3500	11 km	0.08	9 km	0.06
mare-crust only	3220	10 km	0.07	8 km	0.05
crust-mantle only	3220	16 km	0.12	18 km	0.12

# Experimental realization of universal geometric quantum gates with solid-state spins

C. Zu<sup>1</sup>, W.-B. Wang<sup>1</sup>, L. He<sup>1</sup>, W.-G. Zhang<sup>1</sup>, C.-Y. Dai<sup>1</sup>, F. Wang<sup>1</sup> & L.-M. Duan<sup>1,2</sup>

Experimental realization of a universal set of quantum logic gates is the central requirement for the implementation of a quantum computer. In an ‘all-geometric’ approach to quantum computation<sup>1,2</sup>, the quantum gates are implemented using Berry phases<sup>3</sup> and their non-Abelian extensions, holonomies<sup>4</sup>, from geometric transformation of quantum states in the Hilbert space<sup>5</sup>. Apart from its fundamental interest and rich mathematical structure, the geometric approach has some built-in noise-resilience features<sup>1,2,6,7</sup>. On the experimental side, geometric phases and holonomies have been observed in thermal ensembles of liquid molecules using nuclear magnetic resonance<sup>8,9</sup>; however, such systems are known to be non-scalable for the purposes of quantum computing<sup>10</sup>. There are proposals to implement geometric quantum computation in scalable experimental platforms such as trapped ions<sup>11</sup>, superconducting quantum bits<sup>12</sup> and quantum dots<sup>13</sup>, and a recent experiment has realized geometric single-bit gates in a superconducting system<sup>14</sup>. Here we report the experimental realization of a universal set of geometric quantum gates using the solid-state spins of diamond nitrogen–vacancy centres. These diamond defects provide a scalable experimental platform<sup>15–17</sup> with the potential for room-temperature quantum computing<sup>16–19</sup>, which has attracted strong interest in recent years<sup>20</sup>. Our experiment shows that all-geometric and potentially robust quantum computation can be realized with solid-state spin quantum bits, making use of recent advances in the coherent control of this system<sup>15–20</sup>.

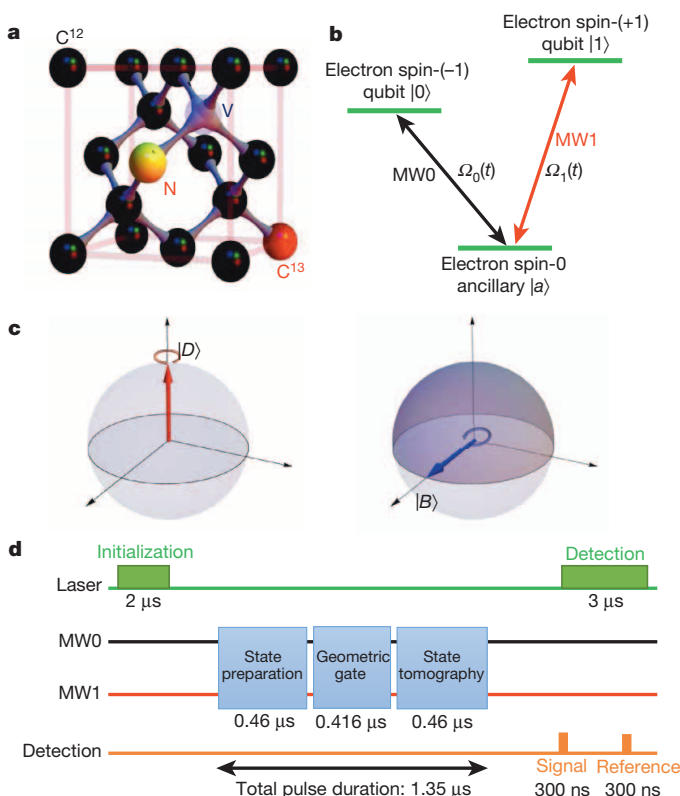
Under adiabatic cyclic evolution, a non-degenerate eigenstate of a quantum system acquires a phase factor, which has a dynamical component proportional to the time integral of the eigenstate energy and a geometric component determined by the global property of the evolution path. This geometric phase, first discovered by Berry<sup>3</sup>, has been linked with many important physics phenomena<sup>21</sup>. If the system has degenerate eigenstates, the Berry phase is replaced by a geometric unitary operator acting on the degenerate subspace, called a holonomy by analogy with differential geometry. The holonomies do not in general commute with each other. In the proposed geometric approach to quantum computation<sup>1,2</sup>, such holonomies are exploited to realize a universal set of quantum gates, compositions of which then can be used to perform arbitrary quantum computation tasks. Because holonomies are determined by global geometric properties, geometric computation is more robust to certain control errors<sup>1,2,6,7</sup>. The implementation of geometric quantum computation has been proposed in several quantum bit (qubit) systems<sup>11–13</sup>; however, it remains experimentally challenging to realize a universal set of gates using holonomies alone, because of the requirements of slow adiabatic evolution and a complicated level structure.

In the recent proposal of non-adiabatic geometric quantum computation<sup>6,22</sup>, universal quantum gates are constructed fully by geometric means without the requirement of adiabatic evolution, thereby combining speed with universality. Under a cyclic evolution of the system Hamiltonian  $H(t)$  (with  $H(\tau) = H(0)$ , where  $\tau$  is the cycle period), we let  $|\xi_l(t)\rangle$  ( $l = 1, 2, \dots, M$ ) denote instantaneous orthonormal bases (moving frames) which coincide with the basis vectors  $|\xi_l\rangle$  of the computational space  $C$  at  $t = 0$  and  $t = \tau$ , with  $|\xi_l(t)\rangle = |\xi_l(0)\rangle = |\xi_l\rangle$ . The evolution operator  $U(\tau)$  for the basis states  $|\xi_l\rangle$  has two contributions: a dynamic

part and a fully geometric part<sup>6</sup>. If the parallel-transport condition  $\langle \xi_l(t) | H(t) | \xi_{l'}(t) \rangle = 0$  is satisfied for any  $l$  and  $l'$  at any time  $t$ , then the dynamic contribution becomes identically zero and  $U(\tau)$  is given by

$$U(\tau) = T \exp \left[ i \int_0^\tau A dt \right] \quad (1)$$

where  $T$  indicates time-ordered integration and  $A = [A_{ll'}] = [\langle \xi_l(t) | i\partial_t | \xi_{l'}(t) \rangle]$  is the  $M \times M$  connection matrix<sup>6</sup>. The form of  $U(\tau)$  is identical to the Wilczek–Zee holonomy in the adiabatic case<sup>4,6</sup>.

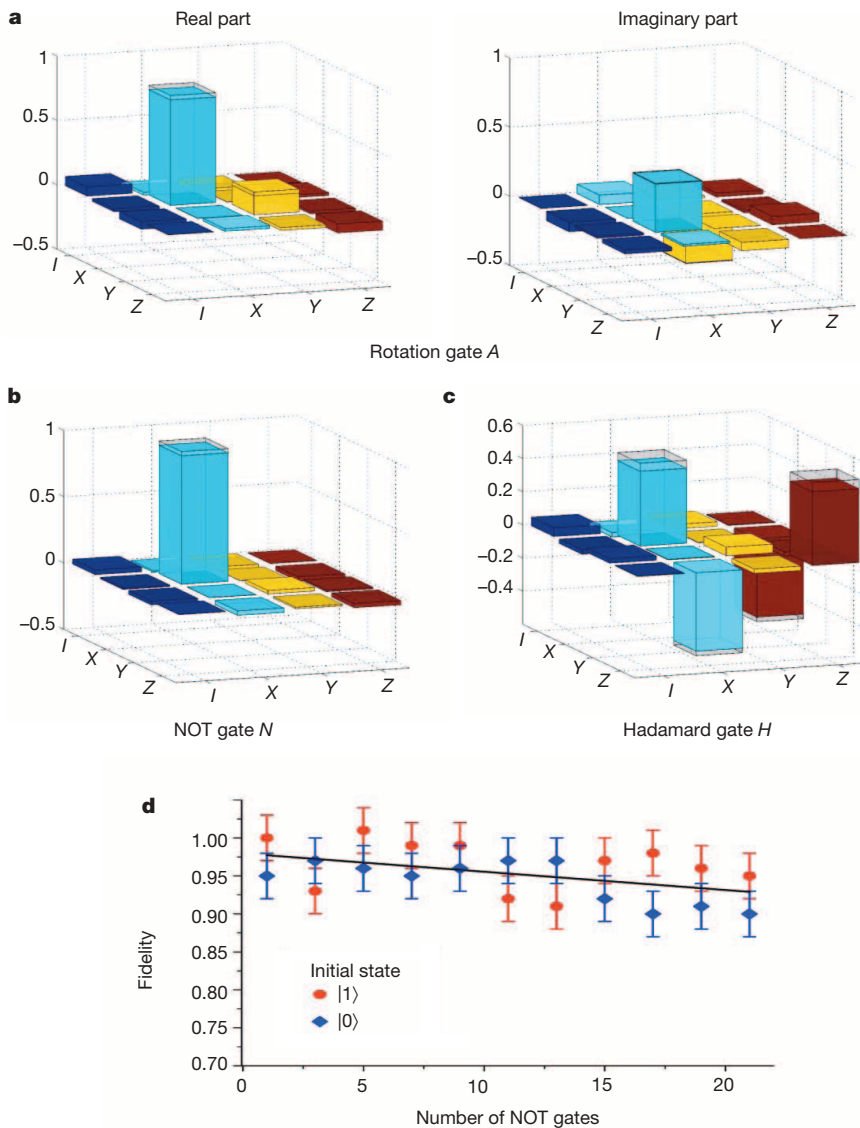


**Figure 1 | Geometric gates in a diamond nitrogen–vacancy centre.**

**a**, Illustration of a nitrogen–vacancy (NV) centre in a diamond with a proximal  $C^{13}$  atom. **b**, Encoding of a qubit in the spin-triplet ground state of the nitrogen–vacancy centre and the microwave coupling configuration. The spin-0 state provides an ancillary level  $|a\rangle$  for geometric manipulation of the qubit. **c**, A geometric picture of the holonomic gates. Under a cyclic Hamiltonian evolution, the dark state  $|D\rangle$  and the bright state  $|B\rangle$  rotate by  $2\pi$  around the North Pole of the Bloch sphere and, respectively, along its equator, acquiring geometric phase of 0 and, respectively,  $\pi$  (half of the swept solid angle). When we choose different forms of the dark and bright states, by controlling parameters in the Hamiltonian, this state-dependent geometric phase leads to the corresponding holonomic gates. **d**, The time sequence for implementation and verification of single-qubit geometric gates.

<sup>1</sup>Center for Quantum Information, IIS, Tsinghua University, Beijing 100084, China. <sup>2</sup>Department of Physics, University of Michigan, Ann Arbor, Michigan 48109, USA.





**Figure 2 | Experimental results for single-bit geometric gates.** **a–c**, The measured process matrix elements for the rotation gate *A* (**a**), the NOT gate *N* (**b**) and the Hadamard gate *H* (**c**). The measured tiny imaginary parts of the process matrices for the NOT and Hadamard gates are not shown. The hollow caps in these figures denote the corresponding matrix elements for the ideal gates. **d**, The measured fidelities of the final states compared with the ideal output (error bars denote s.d.) after application of a sequence of the geometric NOT gates to initial states  $|0\rangle$  and  $|1\rangle$ . By fitting the data under the assumption of independent error for each gate, we obtain the error induced by each NOT gate at  $(0.24 \pm 0.06)\%$ .

Our experiment realizes a universal set of quantum gates using only non-adiabatic holonomies<sup>6</sup>. Single-bit gates, together with entangling controlled-NOT (CNOT) operation, are universal for quantum computation. Our realization is based on the control of electron and nuclear spins in a diamond nitrogen–vacancy centre that effectively form a quantum register<sup>20</sup>. To realize the single-bit geometric gates, we manipulate the electron spin states of a nitrogen–vacancy centre (Fig. 1a) in a synthetic diamond at room temperature (see Methods for a description of the experimental set-up). The nitrogen–vacancy centre has a spin-triplet ground state. We take the Zeeman components  $|m = -1\rangle \equiv |0\rangle$  and  $|m = +1\rangle \equiv |1\rangle$  as the qubit basis states and use  $|m = 0\rangle \equiv |a\rangle$  as an ancillary level for geometric manipulation of the qubit. The spin state is initialized through optical pumping to the  $|m = 0\rangle$  level and read out by distinguishing different fluorescence levels of the states under illumination of a short green laser pulse<sup>20</sup> (see Methods for the calibration of fluorescence levels of different states). We apply a magnetic field of 451 G along the nitrogen–vacancy axis using a permanent magnet. Under this field, the nearby nuclear spins are polarized by optical pumping<sup>23</sup>, enhancing the coherence time of the electron spin.

The transitions from the qubit states  $|0\rangle$  and  $|1\rangle$  to the ancillary level  $|a\rangle$  are coupled by microwave pulses controlled using an arbitrary-waveform generator, with Rabi frequencies  $\Omega_0(t)$  (for the  $|0\rangle \rightarrow |a\rangle$  transition) and  $\Omega_1(t)$  (for the  $|1\rangle \rightarrow |a\rangle$  transition) (Fig. 1b). We vary the amplitude  $\Omega(t) = \sqrt{\Omega_0^2 + \Omega_1^2}$  but fix the ratio  $\Omega_1/\Omega_0 = e^{i\phi} \tan(\theta)$  to be

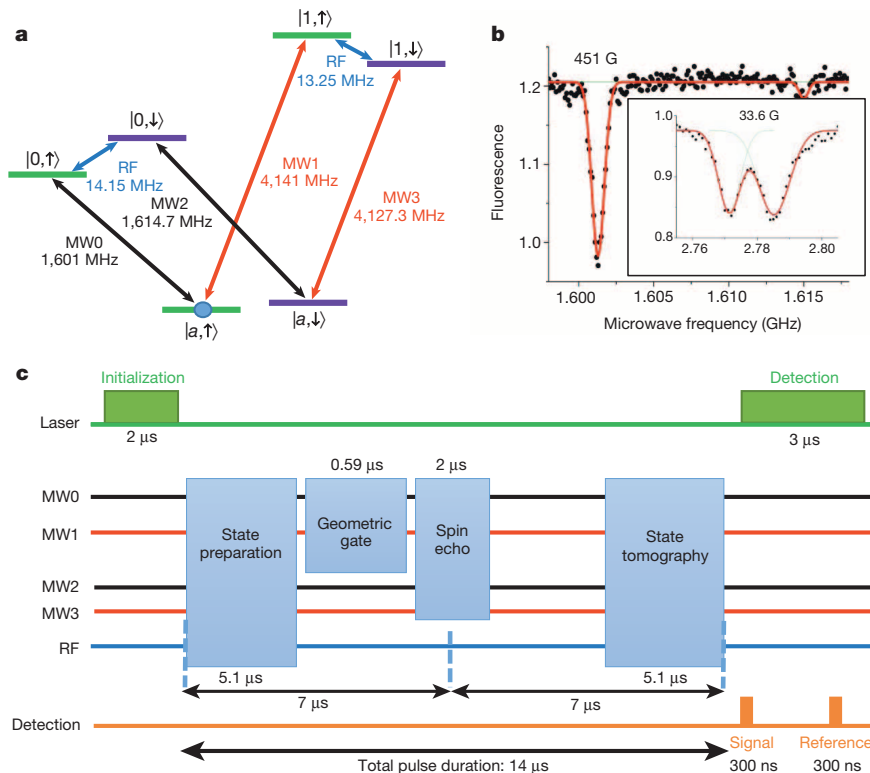
constant. The Hamiltonian for the coupling between these three levels takes the form

$$H_I(t) = \hbar\Omega(t) [(\cos(\theta)|0\rangle + e^{i\phi}\sin(\theta)|1\rangle)\langle a| + \text{H.c.}]$$

where  $\hbar$  is Planck's constant divided by  $2\pi$  and H.c. denotes the Hermitian conjugate. We define the bright state as  $|B\rangle = \cos(\theta)|0\rangle + e^{i\phi}\sin(\theta)|1\rangle$  and the dark state as  $|D\rangle = -e^{-i\phi}\sin(\theta)|0\rangle + \cos(\theta)|1\rangle$ . When  $\Omega(t)$  makes a cyclic evolution with  $\Omega(0) = \Omega(\tau) = 0$ , the bright state evolves as  $|B(t)\rangle = e^{i\alpha(t)}[\cos(\alpha(t))|B\rangle + \sin(\alpha(t))|a\rangle]$ , where  $\alpha(t) \equiv \int_0^t \Omega(t') dt'$ , while the dark state remains unchanged. After a cyclic evolution with  $\alpha(\tau) = \pi$ , the bright and dark states pick up geometric phases of  $\pi$  and 0, respectively (Fig. 1c). We take the moving frame as  $|\xi_0(t)\rangle = \cos(\theta)|B(t)\rangle - e^{i\phi}\sin(\theta)|D\rangle$ ,  $|\xi_1(t)\rangle = e^{-i\phi}\sin(\theta)|B(t)\rangle + \cos(\theta)|D\rangle$ , which makes a cyclic evolution with  $|\xi_i(0)\rangle = |\xi_i(\tau)\rangle = |i\rangle$  ( $i = 0, 1$ ). For this evolution, it can easily be checked that the condition  $\langle \xi_i(t) | H(t) | \xi_j(t) \rangle = 0$  is always satisfied, such that there is no dynamic contribution to the evolution operator  $U(\tau)$  (ref. 6). Using equation (1), we find the holonomy

$$U(\tau) = \begin{bmatrix} -\cos(2\theta) & -e^{i\phi}\sin(2\theta) \\ -e^{-i\phi}\sin(2\theta) & \cos(2\theta) \end{bmatrix}$$

in the computational basis  $\{|0\rangle, |1\rangle\}$ .

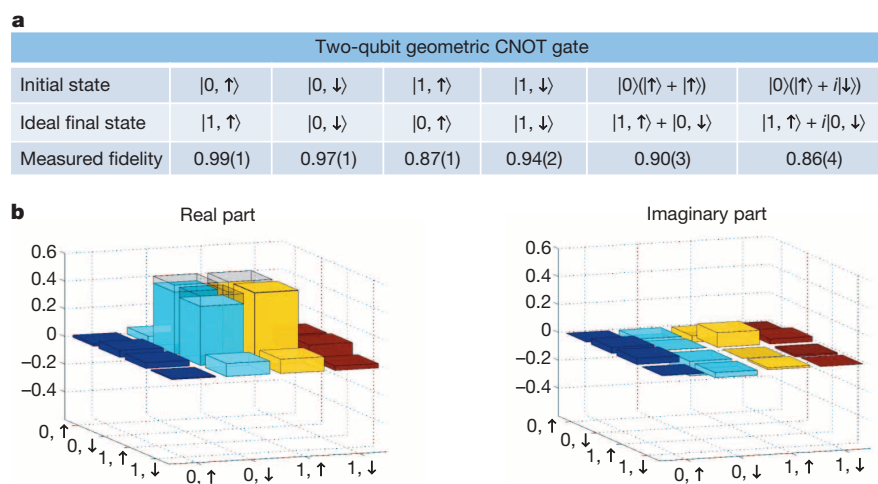


**Figure 3 | Level scheme and pulse sequence for the geometric CNOT gate.** **a**, The level structure of the electron and the nuclear spins for the geometric CNOT gate and the microwave and radio-frequency (RF) coupling configuration. **b**, Optically detected magnetic resonance spectroscopy obtained by measuring the fluorescence level while scanning the frequency of the microwave that couples to the transition between  $|a\rangle$  (spin 0) and  $|1\rangle$  (spin 1). The two dips at 33.6 G (inset) represent the hyperfine splitting caused by the unpolarized nuclear spin. The very asymmetric dips at 451 G indicate that the nuclear spin has been polarized. **c**, The time sequence for implementation

and verification of the geometric CNOT gate between the electron and the nuclear spins. The CNOT gate is implemented by applying the pulses MW0 and MW1 simultaneously. Microwave pulses MW2 and MW3 are used, in addition to MW0 and MW1, to implement a spin echo to increase the spin coherence time. To verify the CNOT gate, we use a combination of MW0–MW3 and a radio-frequency pulse to prepare various initial superposition states and measure the final output in different bases through quantum state tomography.

We evolve the Rabi frequencies  $\Omega_i(t)$  along three different loops, with the parameters  $(\theta, \varphi)$  chosen respectively as  $(3\pi/4, 0)$ ,  $(3\pi/4, \pi/8)$  and  $(5\pi/8, 0)$ . The three geometric gates resulting from these cyclic evolutions are denoted by the NOT gate  $N$ , the rotation gate  $A$  and the Hadamard gate  $H$ , respectively. The combination of gates  $N$  and  $A$  gives the well-known  $\pi/8$ -gate  $T = NA$ . Together,  $N$ ,  $A$  and  $H$  make a universal set of single-bit gates. To characterize these geometric gates, we use quantum process tomography by preparing and measuring the qubit in different

bases<sup>24</sup>, with the time sequence shown in Fig. 1d. The matrix elements for each process are shown in Fig. 2a–c, which are shown, for comparison, with the corresponding elements of the ideal gates. From the process tomography (Methods), we find the process fidelities  $F_p = (96.5 \pm 1.9)\%$ ,  $(96.9 \pm 1.5)\%$  and  $(92.1 \pm 1.8)\%$  for the  $N$ ,  $A$  and  $H$  gates, respectively. The major contribution to the infidelity actually comes from the state preparation and detection error in the quantum process tomography. To measure the intrinsic gate error, we concatenate a series of



**Figure 4 | Experimental results for the geometric CNOT gate.** **a**, Measured output state fidelities of the geometric CNOT gate under a few typical input states, where the number in the bracket represents the error bar (s.d.) in the last digit. **b**, The matrix elements of the output density operator reconstructed through quantum state tomography when the geometric CNOT is applied to the product state  $|0\rangle(|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}$ . The hollow caps denote the matrix elements for the ideal output state under a perfect gate.

gates and examine the fidelity decay as the number of gates increases<sup>19</sup>. As an example, we show in Fig. 2d the fidelity decay by concatenating the NOT gates. From the data, we find that the intrinsic error per gate is about 0.24%. This can be compared with the 1% error rate for the dynamic NOT gate using optimized pulses and the same method of measurement<sup>19</sup>. The achieved high fidelity indicates that geometric manipulation is indeed resilient to control errors.

To realize the geometric quantum CNOT gate, we use one nearby  $C^{13}$  nuclear spin as the control qubit (with the basis vectors  $|\uparrow\rangle$  and  $|\downarrow\rangle$ ) and the nitrogen–vacancy centre electron spin as the target qubit<sup>25</sup>. Both the electron spin and the nuclear spin are polarized through optical pumping under the 451 G magnetic field, which is confirmed by optically detected magnetic resonance spectroscopy (Fig. 3b). The spins are interacting with each other through hyperfine and dipole couplings, and the resultant level configuration is shown in Fig. 3a. By applying state-selective microwave and radio-frequency pulses, we can couple different levels. In particular, for the microwave pulses MW0 and MW1, with respective Rabi frequencies  $\Omega_0(t)$  and  $\Omega_1(t)$ , we have the following coupling Hamiltonian:

$$H_2(t) = \hbar\Omega(t)[(|0,\uparrow\rangle - |1,\uparrow\rangle)\langle a,\uparrow| + \text{H.c.}]/\sqrt{2}$$

Here we have fixed the ratio  $\Omega_1/\Omega_0 = -1$ . Under cyclic evolution of  $\Omega(t)$  with  $\int_0^\tau \Omega(t) dt = \pi$ , we find the holonomy  $U(\tau) = |\uparrow\rangle\langle\uparrow| \otimes N + |\downarrow\rangle\langle\downarrow| \otimes I$  using equation (1), where  $I$  denotes the  $2 \times 2$  unit matrix. This achieves the quantum CNOT gate exactly.

To characterize the geometric CNOT gate, we apply the gate to the qubit basis states as well as their superpositions, and measure the fidelity of the final states relative to the ideal outputs, by quantum state tomography<sup>24</sup>. The superposition of the nuclear spin states required for state preparation and measurement is generated using radio-frequency pulses, which takes longer than it would with microwave pulses owing to the much smaller magnetic moment of the nuclear spin. The electron spin decoherence is significant during the slow radio-frequency pulses. To correct that, we apply a Hahn spin echo in the middle of the whole operation with the time sequence shown in Fig. 3c. The measured state fidelities are listed in Fig. 4a under typical input states. A hallmark of the entangling operation is that the geometric CNOT gate generates entanglement from the initial product state. As an example, for the input state  $|0\rangle \otimes (|\uparrow\rangle + |\downarrow\rangle)$  (unnormalized), the matrix elements of the output density operator are shown in Fig. 4b, with a measured entanglement fidelity of  $(90.2 \pm 2.5)\%$  and a concurrence of  $0.85 \pm 0.05$ , which unambiguously confirms entanglement<sup>10</sup>.

Our experimental realization of a universal set of holonomic gates using individual spins paves the way for all-geometric quantum computation in a solid-state system. The electron and nuclear spins of different nitrogen–vacancy centres can be wired up quantum mechanically to form a scalable network of qubits through, for example, direct dipole interaction<sup>16,18</sup>, spin-chain assisted coupling by the nitrogen dopants<sup>17,26</sup> or photon-mediated coupling<sup>15,27,28</sup>. The technique used here for the geometric realization of universal gates may also find applications in other scalable experimental systems, such as trapped ions or superconducting qubits. The geometric phase is closely related to the topological phase<sup>29,30</sup>, and the demonstration of gates using only holonomies is a step towards realization of topological computation<sup>30</sup>, the most robust way of quantum computing.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 May; accepted 31 July 2014.

- Zanardi, P. & Rasetti, M. Holonomic quantum computation. *Phys. Lett. A* **264**, 94–99 (1999).
- Pachos, J., Zanardi, P. & Rasetti, M. Non-Abelian Berry connections for quantum computation. *Phys. Rev. A* **61**, 010305(R) (2000).
- Berry, M. V. Quantal phase-factors accompanying adiabatic changes. *Proc. R. Soc. Lond. A* **392**, 45–57 (1984).
- Wilczek, F. & Zee, A. Appearance of gauge structure in simple dynamical systems. *Phys. Rev. Lett.* **52**, 2111–2114 (1984).
- Lloyd, S. Computation from geometry. *Science* **292**, 1669 (2001).
- Sjöqvist, E. et al. Non-adiabatic holonomic quantum computation. *New J. Phys.* **14**, 103035 (2012).
- Johansson, M. et al. Robustness of non-adiabatic holonomic gates. *Phys. Rev. A* **86**, 062322 (2012).
- Jones, J. A., Vedral, V., Ekert, A. & Castagnoli, G. Geometric quantum computation using nuclear magnetic resonance. *Nature* **403**, 869–871 (2000).
- Feng, G., Xu, G. & Long, G. Experimental realization of nonadiabatic holonomic quantum computation. *Phys. Rev. Lett.* **110**, 190501 (2013).
- Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2010).
- Duan, L. M., Cirac, J. I. & Zoller, P. Geometric manipulation of trapped ions for quantum computation. *Science* **292**, 1695–1697 (2001).
- Falci, G. et al. Detection of geometric phases in superconducting nanocircuits. *Nature* **407**, 355–358 (2000).
- Solinas, P., Zanardi, P., Zanghi, N. & Rossi, F. Holonomic quantum gates: a semiconductor-based implementation. *Phys. Rev. A* **67**, 062315 (2003).
- Abdumalikov, A. A. et al. Experimental realization of non-Abelian non-adiabatic geometric gates. *Nature* **496**, 482–485 (2013).
- Pfaff, W. et al. Unconditional quantum teleportation between distant solid-state quantum bits. *Science* **345**, 532–535 (2014).
- Neumann, P. et al. Scalable quantum register based on coupled electron spins in a room temperature solid. *Nature Phys.* **6**, 249–253 (2010).
- Yao, N. Y. et al. Scalable architecture for a room temperature solid-state quantum information processor. *Nature Commun.* **3**, 800 (2012).
- Dolde, F. et al. Room-temperature entanglement between single defect spins in diamond. *Nature Phys.* **9**, 139–143 (2013).
- Dolde, F. et al. High-fidelity spin entanglement using optimal control. *Nature Commun.* **5**, 3371 (2014).
- Doherty, M. W. et al. The nitrogen-vacancy colour centre in diamond. *Phys. Rep.* **528**, 1–45 (2013).
- Shapere, A. & Wilczek, F. *Geometric Phases in Physics* (World Scientific, 1989).
- Zhu, S. L. & Wang, Z. D. Implementation of universal quantum gates based on nonadiabatic geometric phases. *Phys. Rev. Lett.* **89**, 097902 (2002).
- Jacques, V. et al. Dynamic polarization of single nuclear spins by optical pumping of nitrogen-vacancy color centers in diamond at room temperature. *Phys. Rev. Lett.* **102**, 057403 (2009).
- White, A. G. et al. Measuring two-qubit gates. *J. Opt. Soc. Am. B* **24**, 172–183 (2007).
- Jelezko, F. et al. Observation of coherent oscillation of a single nuclear spin and realization of a two-qubit conditional quantum gate. *Phys. Rev. Lett.* **93**, 130501 (2004).
- Yao, N. Y. et al. Robust quantum state transfer in random unpolarized spin chains. *Phys. Rev. Lett.* **106**, 040505 (2011).
- Jiang, L., Taylor, J. M., Sorensen, A. S. & Lukin, M. D. Distributed quantum computation based on small quantum registers. *Phys. Rev. A* **76**, 062323 (2007).
- Duan, L. M. & Monroe, C. Quantum networks with trapped ions. *Rev. Mod. Phys.* **82**, 1209 (2010).
- Loredo, J. C., Broome, M. A., Smith, D. H. & White, A. G. Observation of entanglement-dependent two-particle holonomic phase. *Phys. Rev. Lett.* **112**, 143603 (2014).
- Pachos, J. K. *Introduction to Topological Quantum Computation* (Cambridge Univ. Press, 2012).

**Acknowledgements** We thank M. Lukin's group for discussions. This work was supported by the National Basic Research Program of China 2011CBA00302 and the quantum information project from the Ministry of Education of China. In addition, L.-M.D. acknowledges support from the IARPA MUSIQ program, the AFOSR and the ARO MURI program.

**Author Contributions** L.-M.D. had the idea for the experiment and supervised the project. C.Z., W.-B.W., L.H., W.-G.Z., C.-Y.D., F.W. carried out the experiment. L.-M.D. and C.Z. wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.-M.D. (lmduan@umich.edu).



## METHODS

**Experimental set-up.** We use a home-built confocal microscopy, with an oil-immersed objective lens ( $NA = 1.49$ ), to address and detect single nitrogen–vacancy centres in a type-IIa, single-crystal synthetic diamond sample (Element Six). A 532 nm diode laser, controlled by an acoustic optical modulator (AOM), is used for spin-state initialization and detection. We collect fluorescence photons (wavelength ranging from 637 to 850 nm) into a single-mode fibre and detect them using the single-photon counting module (SPCM), with a counting rate of 105 kHz and a signal-to-noise ratio of 15:1. The diamond sample is mounted on a three-axis, closed-loop piezoelectric actuator for submicrometre-resolution scanning. An impedance-matched gold coplanar waveguide (CPW) with a 70  $\mu\text{m}$  gap, deposited on a coverslip, is used for delivery of radio-frequency and microwave signals to the nitrogen–vacancy centre.

In our experiment, we find a single nitrogen–vacancy centre with a proximal  $\text{C}^{13}$  of 13.7 MHz hyperfine strength (Fig. 1). To polarize the nearby nuclear spins ( $\text{C}^{13}$  and the host  $\text{N}^{14}$ ), we apply a magnetic field of 451 G along the nitrogen–vacancy axis using a permanent magnet. Under this field, the electron spin levels  $|m = 0\rangle$  and  $|m = -1\rangle$  become almost degenerate in the optically excited state (called the esLAC, the electron-spin level anti-crossing<sup>23</sup>), which facilitates electron-spin/nuclear-spin flip-flop process during optical pumping. The spin flip-flop process leads to polarization of the nitrogen nuclear spin on the nitrogen–vacancy site and the nearby  $\text{C}^{13}$  nuclear spins after 2  $\mu\text{s}$  green laser illumination<sup>23</sup>. The Zeeman energy from the 451 G magnetic field shifts the respective energy differences between electron spin states  $|m = 0\rangle$  and  $|-1\rangle$  and  $|0\rangle$  and  $|+1\rangle$  from the zero-field splitting, 2,870 MHz, to 1,601 MHz and 4,141 MHz, and shifts the corresponding nuclear-spin hyperfine splittings for the  $|-1\rangle$  and  $|+1\rangle$  levels from 13.7 MHz to 14.15 MHz and 13.25 MHz. Owing to the large energy difference of the  $m = \pm 1$  levels, we apply two independent microwave sources (Rohde–Schwarz), locked by a 10 MHz reference rubidium clock, to address each transition. To adjust the frequency and phase of the microwave pulses, we mix each microwave output with an arbitrary-waveform generator (AWG; Tektronix; 500 MHz sample rate). Radio-frequency signals for nuclear spin manipulation are generated directly by another analogue channel of the AWG. All the microwave and radio-frequency signals are amplified by independent amplifiers, combined through a home-made circuit, and delivered to the CPW. The digital markers of the AWG are used to control the pulse sequence (including the laser and the SPCM) with a timing resolution of 2 ns.

For each experimental cycle, we start the sequence with 2  $\mu\text{s}$  of laser illumination to polarize the nitrogen–vacancy electron spin and nearby nuclear spins, and end it with a 3  $\mu\text{s}$  laser pulse for spin state detection. We collect signal photons for 300 ns right after the detection laser rises (reaches full intensity), and for another

300 ns for reference 2  $\mu\text{s}$  later. With a photon collection rate of 105 kHz, we have an average of 0.03 photon counts per cycle. To measure each datum, we repeat the experimental cycle at least  $10^6$  times, resulting in a total photon count of  $3 \times 10^4$ . The error bars of our data account for the statistical error associated with the photon counting. To calculate the error bar of each datum, we use Monte Carlo simulation by assuming a Poissonian distribution for the photon counts. For each simulation trial, we calculate the value of each datum. Then, by sampling over all the trails according to the Poissonian distribution, we get the statistics of the data (including their mean value and standard deviation (the error bar)).

**Calibration of fluorescence levels for different states.** Owing to the esLAC that induces spin flip-flop during the detection and the imperfect initial polarization of the electron and nuclear spins, each spin component  $|m, m_n\rangle$  ( $m = 0, \pm 1$ ;  $m_n = \uparrow, \downarrow$ ) may fluorescence at different levels. We note that the spins are predominantly in the state  $|m = 0, m_n = \uparrow\rangle$  after the optical pumping. To calibrate the fluorescence level of each state, we therefore associate the detected fluorescence level right after the optical pumping with the state  $|m = 0, m_n = \uparrow\rangle$ . With microwave or radio-frequency  $\pi$ -pulses (the  $\pi$ -pulses are calibrated through Rabi oscillations), we can make a complete transfer between  $|m = 0, m_n = \uparrow\rangle$  and any other  $|m, m_n\rangle$  spin component. For instance, with a  $\pi$ -pulse between  $|m = 0, m_n = \uparrow\rangle$  and  $|m = 0, m_n = \downarrow\rangle$  right after the optical pumping, we associate the detected fluorescence level with the  $|m = 0, m_n = \downarrow\rangle$  state. In this way, the characteristic fluorescence level of each component  $|m, m_n\rangle$  can be calibrated. With the calibrated fluorescence level for each spin component, we then read out the system state after the geometric gates through quantum state tomography<sup>24</sup>.

**Quantum process tomography.** A quantum process can be described by a completely positive map  $\varepsilon$  acting on an arbitrary initial state  $\rho_i$ , transferring it to  $\rho_f \equiv \varepsilon(\rho_i)$ . In quantum process tomography (QPT), we choose a fixed set of basis operators  $\{E_m\}$  so that the map  $\varepsilon(\rho_i) = \sum_{mn} E_m \rho_i E_n^\dagger \chi_{mn}$  is identified with a process matrix  $\chi_{mn}$ . We experimentally measure this process matrix  $\chi$  by the maximum-likelihood technique<sup>24</sup>. For single-bit QPT, we set the basis operators to be  $I = I, X = \sigma_x, Y = -i\sigma_y, Z = \sigma_z$  and choose the four different initial states  $|0\rangle, |1\rangle, (|0\rangle + |1\rangle)/\sqrt{2}$  and  $(|0\rangle - i|1\rangle)/\sqrt{2}$ . We reconstruct the corresponding final density operators through the standard quantum state tomography and use them to calculate the process matrix  $\chi_e$ . This process matrix  $\chi_e$  is compared with the ideal one  $\chi_{id}$  by calculating the process fidelity  $F_p = \text{Tr}(\chi_e \chi_{id})$ . The process fidelity  $F_p$  also determines the average gate fidelity  $\bar{F}$  according to the formula  $\bar{F} = (dF_p + 1)/(d + 1)$  (ref. 24), where  $\bar{F}$  is defined as the fidelity averaged over all possible input states with equal weight and  $d$  is the dimension of the state space (with  $d = 2$  for a single qubit).

# Evanescent-wave and ambient chiral sensing by signal-reversing cavity ringdown polarimetry

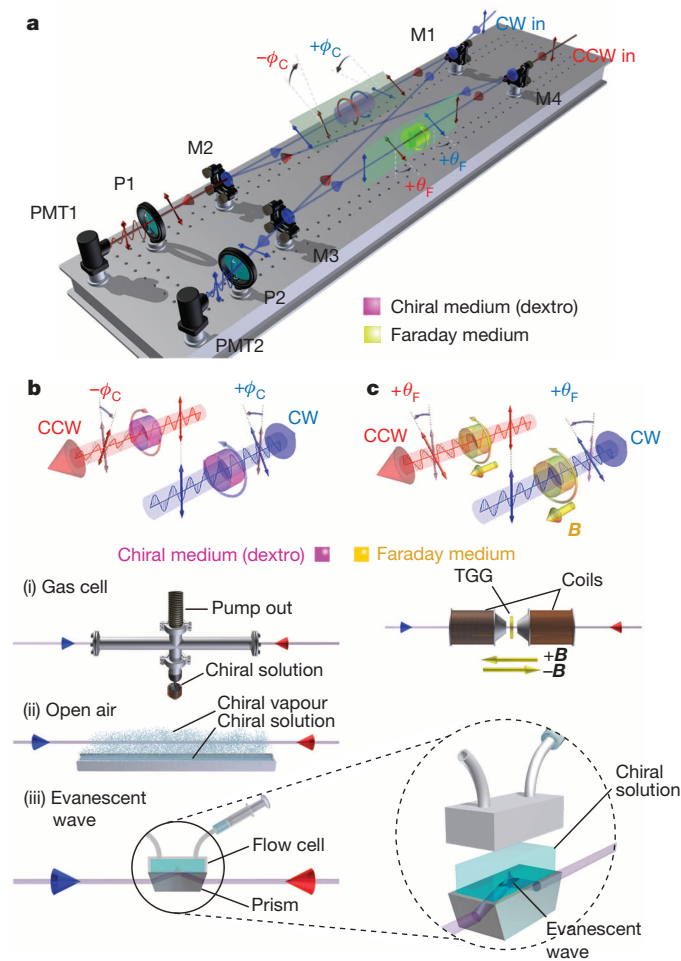
Dimitris Sofikitis<sup>1,2\*</sup>, Lykourgos Bougas<sup>1,2\*</sup>, Georgios E. Katsoprinakis<sup>1,2</sup>, Alexandros K. Spiliotis<sup>1,2</sup>, Benoit Loppinet<sup>1</sup> & T. Peter Rakitzis<sup>1,2</sup>

Detecting and quantifying chirality is important in fields ranging from analytical and biological chemistry to pharmacology<sup>1</sup> and fundamental physics<sup>2</sup>: it can aid drug design and synthesis, contribute to protein structure determination, and help detect parity violation of the weak force. Recent developments employ microwaves<sup>3</sup>, femto-second pulses<sup>4</sup>, superchiral light<sup>5</sup> or photoionization<sup>6</sup> to determine chirality, yet the most widely used methods remain the traditional methods of measuring circular dichroism and optical rotation. However, these signals are typically very weak against larger time-dependent backgrounds<sup>7</sup>. Cavity-enhanced optical methods can be used to amplify weak signals by passing them repeatedly through an optical cavity, and two-mirror cavities achieving up to  $10^5$  cavity passes have enabled absorption and birefringence measurements with record sensitivities<sup>8–10</sup>. But chiral signals cancel when passing back and forth through a cavity, while the ubiquitous spurious linear birefringence background is enhanced. Even when intracavity optics overcome these problems<sup>11–15</sup>, absolute chirality measurements remain difficult and sometimes impossible. Here we use a pulsed-laser bowtie cavity ringdown polarimeter with counter-propagating beams<sup>16,17</sup> to enhance chiral signals by a factor equal to the number of cavity passes (typically  $>10^3$ ); to suppress the effects of linear birefringence by means of a large induced intracavity Faraday rotation; and to effect rapid signal reversals by reversing the Faraday rotation and subtracting signals from the counter-propagating beams. These features allow absolute chiral signal measurements in environments where background subtraction is not feasible: we determine optical rotation from  $\alpha$ -pinene vapour in open air, and from maltodextrin and fructose solutions in the evanescent wave produced by total internal reflection at a prism surface. The limits of the present polarimeter, when using a continuous-wave laser locked to a stable, high-finesse cavity, should match the sensitivity of linear birefringence measurements<sup>8</sup> ( $3 \times 10^{-13}$  radians), which is several orders of magnitude more sensitive than current chiral detection limits<sup>7,14,15</sup> and is expected to transform chiral sensing in many fields.

Our approach to the absolute measurement of chirality makes use of the development of a polarimeter based on a bowtie ring cavity, as proposed recently<sup>16,17</sup>. Unlike a two-mirror cavity, a ring cavity can support two distinct, counter-propagating laser beams (Fig. 1), which we describe as ‘clockwise’ and ‘counter-clockwise’. The symmetry between the beams is broken by a longitudinal magnetic field  $B$  applied to an intracavity magneto-optic window, which induces a Faraday rotation  $\theta_F$ . A chiral sample is introduced to only one arm of the cavity. Faraday and chiral optical rotation have different symmetries:  $\theta_F$  is determined only by  $B$ , and has the same sign for both beams, whereas  $\phi_C$  is determined only by the propagation direction and has opposite signs for the two beams (definitions are given for the laboratory frame; Fig. 1). Therefore, the total single-pass optical rotations for the clockwise and counter-clockwise beams are given by the sum and the difference of  $\theta_F$  and  $\phi_C$ , respectively:  $\theta_{CW} = \theta_F + \phi_C$  and  $\theta_{CCW} = \theta_F - \phi_C$ .

As the beams traverse the cavity, their polarizations rotate with angular frequencies  $\omega_{CW}(\pm B) = (\pm\theta_F + \phi_C)/L$  and  $\omega_{CCW}(\pm B) = (\pm\theta_F - \phi_C)/L$ ,

where the dependence of  $\omega_{CW}$  and  $\omega_{CCW}$  on the direction of the axial magnetic field  $B$  is shown schematically ( $+B$  and  $-B$  refer to opposite directions of the magnetic field along one of the arms of the cavity, with  $B = |B|$ ), and  $L$  is the total round-trip cavity length. The difference  $\Delta\omega(\pm B) = |\omega_{CW}(\pm B)| - |\omega_{CCW}(\pm B)|$  equals  $\pm 2\phi_C/(c/L)$ . This key result



**Figure 1 | Cavity-enhanced polarimeter for chiral sensing.** **a**, The layout of four mirrors (M1–M4), polarizers (P1, P2), photomultiplier tubes (PMT1, PMT2), the chiral sample, the Faraday medium and the counter-propagating laser beams (CW and CCW). **b**, The chiral sample gives opposite laboratory-frame optical rotations for CW and CCW ( $\phi_C^{CW} = -\phi_C^{CCW}$ ). Measurements were performed on chiral samples in a gas cell, in open air and in an evanescent wave at a prism surface. **c**, The Faraday rotator gives the same optical rotation for CW and CCW ( $\theta_F^{CW} = \theta_F^{CCW}$ ), controlled by the magnitude and sign of the applied magnetic field ( $B$ ). TGG, terbium gallium garnet crystal.

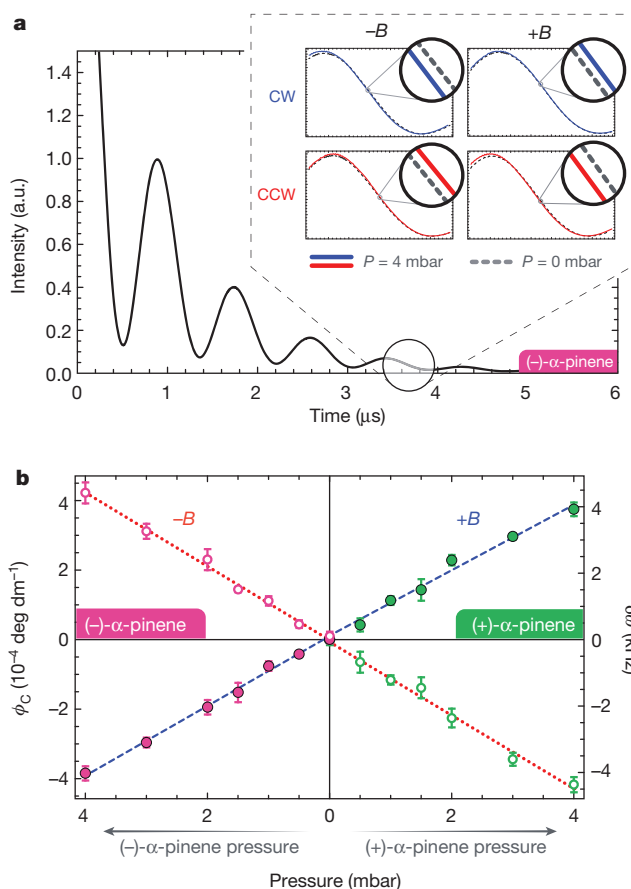
<sup>1</sup>Institute of Electronic Structure and Laser, Foundation for Research and Technology - Hellas, 71110 Heraklion, Greece. <sup>2</sup>Department of Physics, University of Crete, 71003 Heraklion, Greece.

\*These authors contributed equally to this work.

shows that reversing the sign of  $B$  inverts the sign of the measured angle  $\phi_C$  (Fig. 2). Using this signal reversal yields  $\Delta\omega(+B) - \Delta\omega(-B) = 4\phi_C(c/L)$ . For each subtraction, the chiral signal  $\phi_C(c/L)$ , which is odd under reversal of the light propagation direction or of  $B$ , doubles. In contrast, all background signals, which are even under either reversal, cancel.

These signal reversals are demonstrated in three different environments: (1) pressure-controlled chiral vapours, (2) open-air chiral vapours and (3) chiral solutions at a prism surface probed using evanescent waves. Experiment (1) demonstrates the full symmetry of the signal reversals in the absence of large noise, experiments (2) and (3) take advantage of these signal reversals to measure chiral optical rotation in high-noise environments.

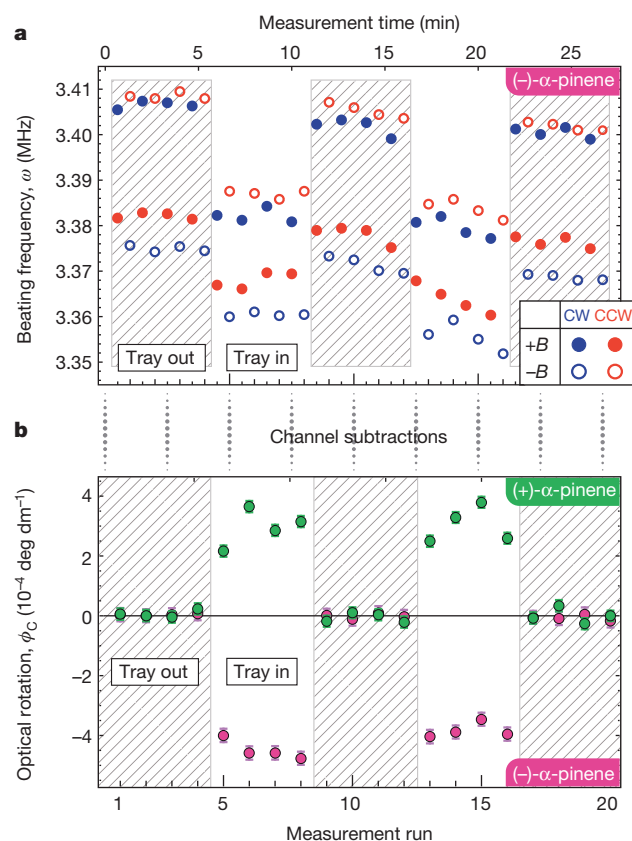
For experiment (1), various pressures of (+)- and (-)- $\alpha$ -pinene are introduced into an intracavity cell, and the four polarization frequencies  $\omega_{CW}^{+B}$ ,  $\omega_{CW}^{-B}$ ,  $\omega_{CCW}^{+B}$  and  $\omega_{CCW}^{-B}$  are measured (Fig. 2a). The angle  $\phi_C$  is determined for both (+)- and (-)-enantiomers, and for both  $+B$  and  $-B$ , and is plotted versus pressure in Fig. 2b. The optical rotation  $\phi_C$  varies linearly with pressure, and the expected symmetry is obtained:  $\phi_C$  reverses sign for each inversion of enantiomer or  $B$ . Notice how an inversion of  $B$  allows the determination of absolute optical rotation, without needing to change the gas pressure. We determined the specific rotation at 800 nm for the (+)- and (-)-enantiomers to be  $+17.57 \pm 0.57$  and  $-18.04 \pm 0.98^\circ \text{ dm}^{-1} \text{ g}^{-1} \text{ ml}$ , respectively, with sensitivity similar to that in ref. 15.



**Figure 2 | Gas-cell optical rotation.** **a**, Experimental signals showing the polarization oscillation frequencies  $\omega_{CW}^{+B}$ ,  $\omega_{CW}^{-B}$ ,  $\omega_{CCW}^{+B}$  and  $\omega_{CCW}^{-B}$  in the exponential decay, for 4 mbar and 0 mbar of (-)- $\alpha$ -pinene vapour. Notice the sign change of the frequency shift between CW and CCW or  $+B$  and  $-B$  signal pairs (insets). **b**, Measurements of optical rotations  $\phi_C$  and frequency shift difference  $\delta\omega$ , for  $+B$  and  $-B$  and for (+)- and (-)- $\alpha$ -pinene vapours, as functions of gas pressure. Each data point is the average of four data sets of difference signals ( $\omega_{CW}^{+B} - \omega_{CCW}^{+B}$ ), each of 8,000 laser shots. Error bars are  $2\sigma$  confidence intervals.

For experiment (2), we perform open-air measurements by inserting and removing a tray filled with liquid (+)- or (-)- $\alpha$ -pinene, below one of the arms of the cavity, and measuring the optical rotation of the vapour that evaporates. The four measured frequencies  $\omega_{CW}^{+B}$ ,  $\omega_{CW}^{-B}$ ,  $\omega_{CCW}^{+B}$  and  $\omega_{CCW}^{-B}$  shown in Fig. 3a illustrate that each of the four traces separately yields incorrect results for the optical rotation, some even giving the wrong sign. This is because strong variations in the index of refraction of the vapour alter the alignment of the cavity, yielding spurious changes in the polarization beating frequencies, which are larger than those from the optical rotation. Also, a temperature drift causes a downward slope in all four channels. However, the result of the two signal reversals (Fig. 3b) yields a constant null signal (tray removed) and measurement of optical rotation of the open-air (+)- and (-)- $\alpha$ -pinene vapours. By comparing Fig. 3b with Fig. 2b, we deduce that the partial pressure of the vapours was about 4 mbar, in agreement with the vapour pressure of  $\alpha$ -pinene.

Finally, for experiment (3) on solutions of maltodextrin, fructose and non-chiral glycerol as the reference sample, we insert a Dove prism into the beam such that the laser beams undergo total internal reflection with angle of incidence  $\theta = 84^\circ$  (Figs 1 and 4). Figure 4a shows  $\omega_{CW}^{+B}$ ,  $\omega_{CW}^{-B}$ ,  $\omega_{CCW}^{+B}$  and  $\omega_{CCW}^{-B}$ . Time-dependent birefringent variations in the prism cause drifts in the polarization oscillations, which mask chiral optical rotation signals in any single trace. Figure 4b shows the chiral optical rotation signal obtained from the two reversals, which now show clear signal differences between the three solutions. Figure 4c shows the dependence of the maltodextrin and fructose signals on the solution refractive index  $n$ , and emphasizes a strong increase as  $n$  approaches  $n_{\text{critical}} = n_p \sin(\theta) = 1.445$



**Figure 3 | Open-air optical rotation.** **a**, The four polarization frequencies  $\omega_{CW}^{+B}$ ,  $\omega_{CW}^{-B}$ ,  $\omega_{CCW}^{+B}$  and  $\omega_{CCW}^{-B}$  are shown for open-air measurements of (-)- $\alpha$ -pinene vapour, evaporating from a tray which is periodically inserted and removed. Each data point is the average of 4,000 laser shots. Each of the four polarization frequencies is dominated by spurious signals and background drifts. **b**, Subtractions of the polarization signals yield the open-air optical rotation, shown for (+)- and (-)- $\alpha$ -pinene vapours and a stable background. The  $1\sigma$  statistical error bars are determined from the nonlinear regression analysis of the averaged ringdown traces.



(where  $n_p = 1.453$  is the refractive index of the prism). An analytical expression is derived for the optical rotation from a chiral sample in an evanescent wave,  $\phi_{EW}$ , according to the Drude–Condon model<sup>18</sup> for Maxwell's equations in isotropic optically active media, using the treatment developed in refs 19, 20:

$$\phi_{EW} \approx \frac{\Delta n}{n} \frac{N}{1 - N^2} \frac{\cos(\theta)}{\sqrt{\sin^2(\theta) - N^2}} \quad (1)$$

Here  $\theta$  is the incidence angle,  $\Delta n = (n_+ - n_-)$ ,  $n = (n_+ + n_-)/2$ ,  $n_+$  and  $n_-$  are respectively the refractive indices of the chiral sample for left- and

right-circularly polarized light, and  $N = (n/n_p)$ . The data agree well with theoretical predictions, which are calculated from equation (1), using  $\Delta n_M = 4.25 \times 10^{-6}$ ,  $c_M = (2.665n - 3.545) \times 10^{-5}$  for maltodextrin and  $\Delta n_F = -2.28 \times 10^{-6}$ ,  $c_F = (2.635 - 1.955n) \times 10^{-5}$  for fructose (determined from single-pass optical rotation measurements through a 10 cm cell), where  $c_M$  and  $c_F$  are the concentrations (in grams per cubic centimetre) of the maltodextrin and fructose solutions, respectively. We note that  $\phi_{EW}$  increases sharply near the critical angle ( $N \approx \sin(\theta)$ ), and even more so as  $N$  approaches 1 (near index matching, when also  $\sin(\theta) \approx 1$ ). We approached index matching closely with  $N = 1.442/1.453 = 0.9924$ .

To better understand the physics behind equation (1), we express  $\phi_{EW}$  in terms of the Goos–Hänchen shift  $L_{GH}$ :  $\phi_{EW} = (\pi/\lambda)[\Delta n/(1 - N^2)]$

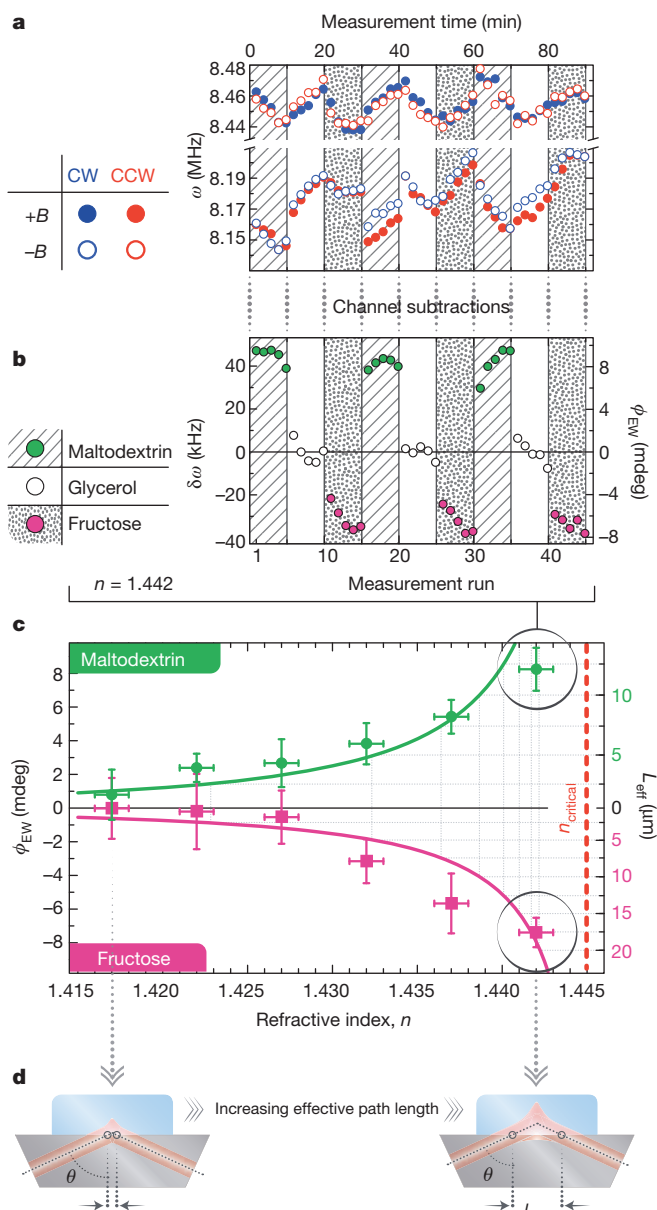
$[\cos^2(\theta)/\sin(\theta)]L_{GH}$ , where  $L_{GH} = \lambda \tan(\theta)/\pi n_p \sqrt{\sin^2(\theta) - N^2}$  is the shift of the light beam at total internal reflection<sup>21</sup> (Fig. 4d), which is the expected relevant length scale for evanescent-wave optical rotation. For an equivalent transmission measurement the optical rotation would be  $\phi_T = (\pi \Delta n/\lambda)L_{eff}$ , where  $L_{eff}$  is the effective path length. Setting  $\phi_{EW} = \phi_T$  yields  $L_{eff} = (\cos^2(\theta)/\sin(\theta)(1 - N^2))L_{GH}$ . Away from the critical angle,  $L_{eff} < L_{GH}$ . Near the critical angle and index matching ( $N \approx \sin(\theta) \approx 1$ ),  $L_{eff} \approx L_{GH}$ , showing that in this case the Goos–Hänchen shift is the effective evanescent-wave optical rotation path length.

The cavity-enhanced polarimetric methods presented here can be extended to a continuous-wave laser locked to a stable, high-finesse cavity, which is much more sensitive<sup>8</sup> but also experimentally more complex. This should increase chiral sensitivities for conventional optical rotation and circular dichroism measurements by several orders of magnitude, and allow the routine analysis of subnanolitre volumes. Applications include the study of surface chirality (for which large effects have been recently shown)<sup>22</sup>, coupling of optical rotation and circular dichroism to gas and liquid chromatography for sensitive chiral analysis, monitoring of protein folding, and measurement of parity violation in atoms and molecules for which insufficient path lengths are otherwise available<sup>16,17</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 May; accepted 9 July 2014.

Published online 10 September 2014.



**Figure 4 | Evanescent-wave optical rotation.** **a**, The four polarization frequencies  $\omega_{CW}^+$ ,  $\omega_{CW}^-$ ,  $\omega_{CCW}^+$  and  $\omega_{CCW}^-$  are shown for evanescent-wave measurements of maltodextrin, glycerol (a non-chiral reference sample with zero optical rotation) and fructose solutions, all with refractive index  $n = 1.442$ . The solutions are changed with a flow cell every 10 min. Each data point is the average of 4,000 laser shots. **b**, Subtractions of the polarization signals give the optical rotations of the samples in the evanescent wave. **c**, Measurements for solutions with  $n = 1.417$ – $1.442$ . Error bars are  $2\sigma$  confidence intervals determined from the 15 points in **b**. The theoretical curves are generated using equation (1). **d**, Goos–Hänchen shift  $L_{GH}$  of the evanescent wave.

1. Berova, N., Polavarapu, P. L., Nakanishi, K. & Woody, W. (eds) *Comprehensive Chiroptical Spectroscopy* Vols 1, 2 (Wiley, 2012).
2. Bouchiat, M. A. & Bouchiat, C. Parity violation in atoms. *Rep. Prog. Phys.* **60**, 1351–1396 (1997).
3. Patterson, D., Schnell, M. & Doyle, J. M. Enantiomer-specific detection of chiral molecules via microwave spectroscopy. *Nature* **497**, 475–477 (2013).
4. Rhee, H. J. *et al.* Femtosecond characterization of vibrational optical activity of chiral molecules. *Nature* **458**, 310–313 (2009).
5. Tang, Y. Q. & Cohen, A. E. Enhanced enantioselectivity in excitation of chiral molecules by superchiral light. *Science* **332**, 333–336 (2011).
6. Janssen, M. H. M. & Powis, I. Detecting chirality in molecules by imaging photoelectron circular dichroism. *Phys. Chem. Chem. Phys.* **16**, 856–871 (2014).
7. Busch, K. W. & Busch, M. A. (eds) *Chiral Analysis* (Elsevier, 2006).
8. Durand, M., Morville, J. & Romanini, D. Shot-noise-limited measurement of sub-parts-per-trillion birefringence phase shift in a high-finesse cavity. *Phys. Rev. A* **82**, 031803 (2010).
9. Berden, G., Peeters, R. & Meijer, G. Cavity ring-down spectroscopy: experimental schemes and applications. *Int. Rev. Phys. Chem.* **19**, 565–607 (2000).
10. Engeln, R., Berden, G., Peeters, R. & Meijer, G. Cavity enhanced absorption and cavity enhanced magnetic rotation spectroscopy. *Rev. Sci. Instrum.* **69**, 3763–3769 (1998).
11. Evtuhov, V. & Siegman, A. E. A. “Twisted-mode” technique for obtaining axially uniform energy density in a laser cavity. *Appl. Opt.* **4**, 142–143 (1965).
12. Kastler, A. Champ lumineux stationnaire à structure hélicoïdale dans une cavité laser. *C. R. Acad. Sci. Paris B* **271**, 999–1001 (1970).
13. Poirson, J., Vallet, M., Bretenaker, F., Le Floch, A. & Thepot, J. Resonant cavity gas-phase polarimeter. *Anal. Chem.* **70**, 4636–4639 (1998).
14. Müller, T., Wiberg, K. B. & Vaccaro, P. H. Cavity ring-down polarimetry (CRDP): a new scheme for probing circular birefringence and circular dichroism in the gas phase. *J. Phys. Chem. A* **104**, 5959–5968 (2000).
15. Wilson, S. M., Wiberg, K. B., Cheeseman, J., Frisch, M. J. & Vaccaro, P. Nonresonant optical activity of isolated organic molecules. *J. Phys. Chem. A* **109**, 11752–11764 (2005).

16. Bougas, B., Katsoprinakis, G. E., von Klitzing, W., Sapirstein, J. & Rakitzis, T. P. Cavity-enhanced parity-nonconserving optical rotation in metastable Xe and Hg. *Phys. Rev. Lett.* **108**, 210801 (2012).
17. Bougas, B., Katsoprinakis, G. E., von Klitzing, W. & Rakitzis, T. P. Fundamentals of cavity-enhanced polarimetry for parity-nonconserving optical rotation measurements: application to Xe, Hg and I. *Phys. Rev. A* **89**, 052127 (2014).
18. Condon, E. U. Theories of optical rotatory power. *Rev. Mod. Phys.* **9**, 432–457 (1937).
19. Silverman, M. P. Reflection and refraction at the surface of a chiral medium: comparison of gyrotropic constitutive relations invariant or noninvariant under a duality transformation. *J. Opt. Soc. Am. A* **3**, 830–837 (1986).
20. Lekner, J. Optical properties of isotropic chiral media. *Pure Appl. Opt.* **5**, 417–443 (1996).
21. Goos, F. & Hänchen, H. Ein neuer und fundamentaler Versuch zur Totalreflexion. *Ann. Phys.* **436**, 333–346 (1947).
22. Emile, J. *et al.* Giant optical activity of sugar in thin soap films. *J. Coll. Inter. Sci.* **408**, 113–116 (2013).

**Acknowledgements** This research was supported by the ERC grant TRICEPS (grant no. 207542), and the FP7 IAPP Programme SOFORT (PIAPGA-2009-251598). B.L. acknowledges the FP7 Infrastructure programme ESMI (CP&CSA-2010-262348) for partial support. We thank P. Tzallas for access to the Attosecond labs at IESL-FORTH.

**Author Contributions** L.B. constructed the experiment, performed the gas-cell and open-air experiments, and analysed the data. D.S. performed the evanescent-wave experiments and analysed the data. G.E.K. and A.K.S. developed the data acquisition and analysis software, and assisted in the experiments. G.E.K. prepared the figures. B.L. derived the evanescent-wave optical rotation equations. T.P.R. had the idea for and directed the experiments, and wrote the manuscript. All authors provided important suggestions for the experiments, discussed the results and contributed to the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.P.R. ([ptr@iesl.forth.gr](mailto:ptr@iesl.forth.gr)).

## METHODS

A 1.3 mJ, 800 nm, 35 fs laser pulse was split and the resulting beams sent in different directions ('clockwise' and 'counterclockwise'). We note that the laser half-width at half-maximum bandwidth of 20 nm did not significantly affect the measurement of chiral optical rotation at the central wavelength of 800 nm for molecular spectra without sharp optical rotation features, such as those studied here. For the study of sharp optical rotation spectra, our cavity ringdown polarimeter is fully compatible with narrow-bandwidth pulsed lasers that are typically used in cavity ringdown spectroscopy<sup>9,10,14,15</sup>. Note that no temperature stabilization or vibration isolation was employed. The mirrors had reflectivity  $R \approx 99.7\%$  and the cavity length was  $L = 3.6$  m. All intracavity optics were antireflection coated for 800 nm, with reflectivities below 0.01% (ATF Boulder). The gas cell and the open-air tray were both of length  $l_0 = 0.75$  m. The  $70^\circ$  fused-silica prism had dimensions  $80 \times 25 \times 25$  mm. The time-dependent intensity of the output light decayed exponentially<sup>13</sup> as  $e^{-t/\tau_0}$ , where the photon lifetime  $\tau_0 = L/c(1 - R^4) \approx 1$   $\mu$ s (Fig. 2a). Faraday rotation  $\theta_F \approx 2.5$ – $4^\circ$  was generated by applying a 0.2–0.3 T magnetic field to a 3 mm-thick terbium gallium garnet crystal. Using a polarizer at the output, the optical rotation appears in the clockwise signal  $I_{CW}$  as an oscillation with frequency  $\omega_{CW}$ ,  $I_{CW} = I_0 e^{-t/\tau_0} [\cos^2(\omega_{CW}t) + \beta]$ , and in the counterclockwise signal with frequency  $\omega_{CCW}$ :  $I_{CCW} =$

$I_0 e^{-t/\tau_0} [\cos^2(\omega_{CCW}t) + \beta]$ , where  $I_0$  is the output intensity at  $t = 0$  (Fig. 2a) and  $\beta$  is a fit parameter that accounts for reduced amplitude modulation<sup>14,15</sup> (caused by imperfections, such as imperfect polarization alignment, birefringence and detector saturation) and is typically less than 0.1. The laser repetition rate was 1 kHz; however, the data acquisition rate was 100 ringdown traces per second, limited by the digital oscilloscope. The magnetic field reversal rate was between 0.02 and 0.045 Hz. The data traces were fitted with the  $I_{CW}$  and  $I_{CCW}$  fit functions, and  $\omega_{CW}$  and  $\omega_{CCW}$  are determined using a nonlinear regression analysis. The magnetic field was reversed between each data point.

Enantiopure (+)- and (–)- $\alpha$ -pinene (Sigma Aldrich) were used. Maltodextrin and fructose were bought commercially. Maltodextrin and fructose solutions (with concentrations between 50 and 60%) and glycerol–water solutions are prepared with refractive indices  $n$  ranging from 1.417 to 1.442 (at 0.005 intervals). For the evanescent-wave set-up (Fig. 1b, (iii)), a magnesium fluoride compensator was inserted to reduce the birefringence  $\delta$  of the prism (typically  $10$ – $20^\circ$ ) to about  $0.5^\circ$  for both beams (the large, position-sensitive birefringence in the prism and imperfect beam alignment precluded better compensation). Modelling the depolarization effects of birefringence on the measurement of  $\omega_{CW}$  and  $\omega_{CCW}$  (refs 16, 17), the ratio  $\delta/\theta_F < 0.2$  yields a correction coefficient  $q^2 > 0.99$ , so that the effect of birefringence is less than 1%.



# Direct observations of evolving subglacial drainage beneath the Greenland Ice Sheet

Lauren C. Andrews<sup>1,2</sup>, Ginny A. Catania<sup>1,2</sup>, Matthew J. Hoffman<sup>3,4</sup>, Jason D. Gulley<sup>1,5</sup>, Martin P. Lüthi<sup>6,7</sup>, Claudia Ryser<sup>7</sup>, Robert L. Hawley<sup>8</sup> & Thomas A. Neumann<sup>4</sup>

Seasonal acceleration of the Greenland Ice Sheet is influenced by the dynamic response of the subglacial hydrologic system to variability in meltwater delivery to the bed<sup>1,2</sup> via crevasses and moulins (vertical conduits connecting supraglacial water to the bed of the ice sheet). As the melt season progresses, the subglacial hydrologic system drains supraglacial meltwater more efficiently<sup>1–4</sup>, decreasing basal water pressure<sup>4</sup> and moderating the ice velocity response to surface melting<sup>1,2</sup>. However, limited direct observations of subglacial water pressure<sup>4–7</sup> mean that the spatiotemporal evolution of the subglacial hydrologic system remains poorly understood. Here we show that ice velocity is well correlated with moulin hydraulic head but is out of phase with that of nearby (0.3–2 kilometres away) boreholes, indicating that moulins connect to an efficient, channelized component of the subglacial hydrologic system, which exerts the primary control on diurnal and multi-day changes in ice velocity. Our simultaneous measurements of moulin and borehole hydraulic head and ice velocity in the Paakitsoq region of western Greenland show that decreasing trends in ice velocity during the latter part of the melt season cannot be explained by changes in the ability of moulin-connected channels to convey supraglacial melt. Instead, these observations suggest that decreasing late-season ice velocity may be caused by changes in connectivity in unchannelized regions of the subglacial hydrologic system. Understanding this spatiotemporal variability in subglacial pressures is increasingly important because melt-season dynamics affect ice velocity beyond the conclusion of the melt season<sup>8–10</sup>.

In the ablation zone of the Greenland Ice Sheet (GIS), moulins deliver surface melt to the base of the ice sheet<sup>11</sup>, where fluctuations in meltwater supply modulate ice motion<sup>2,12,13</sup>. The relationship between surface melting and ice velocity is thought to reflect the structure and evolution of the subglacial hydrologic system<sup>1,2,12,13</sup>. Ice acceleration occurs when meltwater input exceeds the hydraulic capacity of the subglacial system, causing englacial and subglacial water storage and resulting in the widespread reduction in basal friction<sup>12–14</sup>. Subglacial water pressure and ice velocity of the GIS are thought to remain elevated until a channelized drainage system develops, decreasing subglacial water pressure and ice velocity by efficiently routing surface meltwater to the glacier terminus via moulins<sup>1,3,4</sup>. Short-term increases in ice velocity can persist following channelization owing to temporary imbalances between the capacity of this efficient drainage system and the magnitude of surface water input<sup>14,15</sup> from melt<sup>12</sup> and precipitation events<sup>10</sup> or supraglacial lake drainage<sup>13,16</sup>. In this paradigm, channelized systems have been considered the key component governing ice-velocity sensitivity to supraglacial water input. Accordingly, moulin hydraulic head and ice velocity should both decrease seasonally as drainage efficiency increases.

Despite observations of decreasing minimum velocities during much of the melt season<sup>1,9,10,13</sup>, the role of channelization beneath some regions of the GIS may be limited by shallow surface slopes and limited basal conduit meltback<sup>6</sup>. An extensive body of research on alpine glaciers

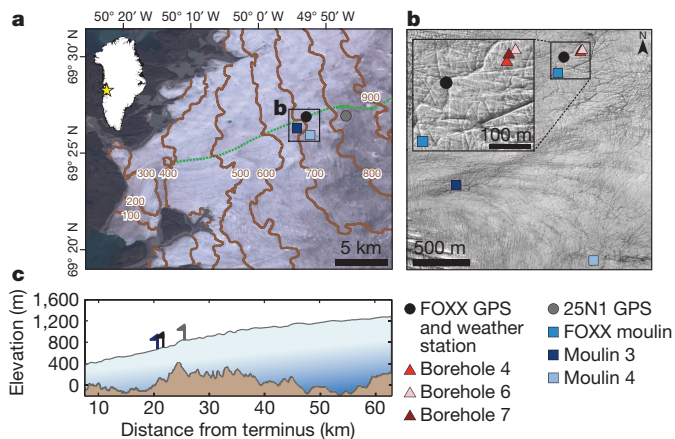
highlights the central role of channelization in subglacial drainage evolution<sup>14,17–19</sup>, but also documents significant complexity in the unchannelized regions of the bed<sup>17–22</sup>. Borehole observations reveal that some portions of the unchannelized region transmit meltwater that is sourced from channels<sup>17,18</sup>, whereas other unchannelized regions of the bed are hydraulically isolated and receive little to no direct water input<sup>17,21,22</sup>. Basal pressure in these isolated regions responds to changes in the active drainage regions through transfer of mechanical support<sup>8,17,22</sup> or as cavity volumes<sup>21</sup> or pore volumes in subglacial sediment fluctuate in response to ice motion. While changing connectivity between active and isolated components of the subglacial drainage system has been directly observed<sup>17,22</sup>, it is not believed to drive seasonal trends in ice velocity on alpine glaciers owing to the dominant control of subglacial channels<sup>17,18,21</sup>.

Here we use borehole hydraulic heads to explore the response of the unchannelized region of the bed to channelized regions. Few direct measurements of water pressure have been made in channelized regions of the GIS bed<sup>4,6</sup>, owing to the limited number of borehole studies<sup>5–7</sup> and their inherently limited spatial extent. Consequently, we also measure water level in three different moulins to constrain hydraulic head in active regions of the subglacial hydrologic system. These borehole and moulin hydraulic heads are coupled with coincident measurements of surface ice velocity, bed separation, and air temperature to characterize relationships between ice dynamics and the subglacial hydrologic system during the 2011 and 2012 melt seasons.

At our primary field site in the ablation zone of Sermeq Avannarleq in western Greenland (Fig. 1; 69° 27' N, 49° 53' W; Extended Data Table 1), we drilled seven boreholes to the bed using a hot water drill and instrumented three of these with pressure transducers in 2011 (ref. 7) (Methods). The ice thickness in our instrumented boreholes was between 614 m and 624 m, and each borehole either drained slowly or did not drain before closure. The combination of gradual drainage following drilling and results from pump tests suggest borehole connection to unchannelized regions of the bed<sup>17,18,22</sup> (Methods). During 2011, we instrumented the FOXX moulin, about 0.3 km southwest of the boreholes, with a pressure transducer. In 2012, we instrumented moulins 3 and 4 with pressure transducers, 1.6 km and 1.9 km from the boreholes, respectively (Methods). Because moulin instrumentation could not proceed until the snowline had retreated past the field sites, these moulin measurements capture relationships between subglacial hydrology and ice motion after channelization is inferred to have begun<sup>4</sup>. We derive ice velocity and bed separation from Global Positioning System (GPS) installations at several locations (Methods; Extended Data Fig. 2). We use meteorological observations from a weather station co-located with the FOXX GPS to determine periods of increasing surface melt (Methods).

The magnitude and phase of moulin and borehole measurements differ substantially in their relationship to ice velocity (Fig. 2; Extended Data Fig. 1), suggesting that each monitored a different component of the subglacial system. Moulin hydraulic heads were highly variable, with

<sup>1</sup>Institute for Geophysics, Jackson School of Geosciences, The University of Texas at Austin, Austin, Texas 78758, USA. <sup>2</sup>Department of Geological Sciences, Jackson School of Geosciences, The University of Texas at Austin, Austin, Texas 78712, USA. <sup>3</sup>Fluid Dynamics and Solid Mechanics Group, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA. <sup>4</sup>NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA. <sup>5</sup>Department of Geological and Mining Engineering and Sciences, Michigan Technological University, Houghton, Michigan 49931, USA. <sup>6</sup>Glaciology and Geomorphodynamics Group, Physical Geography Division, Department of Geography, University of Zürich, 8057 Zürich, Switzerland. <sup>7</sup>Laboratory of Hydraulics, Hydrology and Glaciology, Swiss Federal Institute of Technology (ETH) Zürich, 8093 Zürich, Switzerland. <sup>8</sup>Department of Earth Sciences, Dartmouth College, Hanover, New Hampshire 03755, USA.



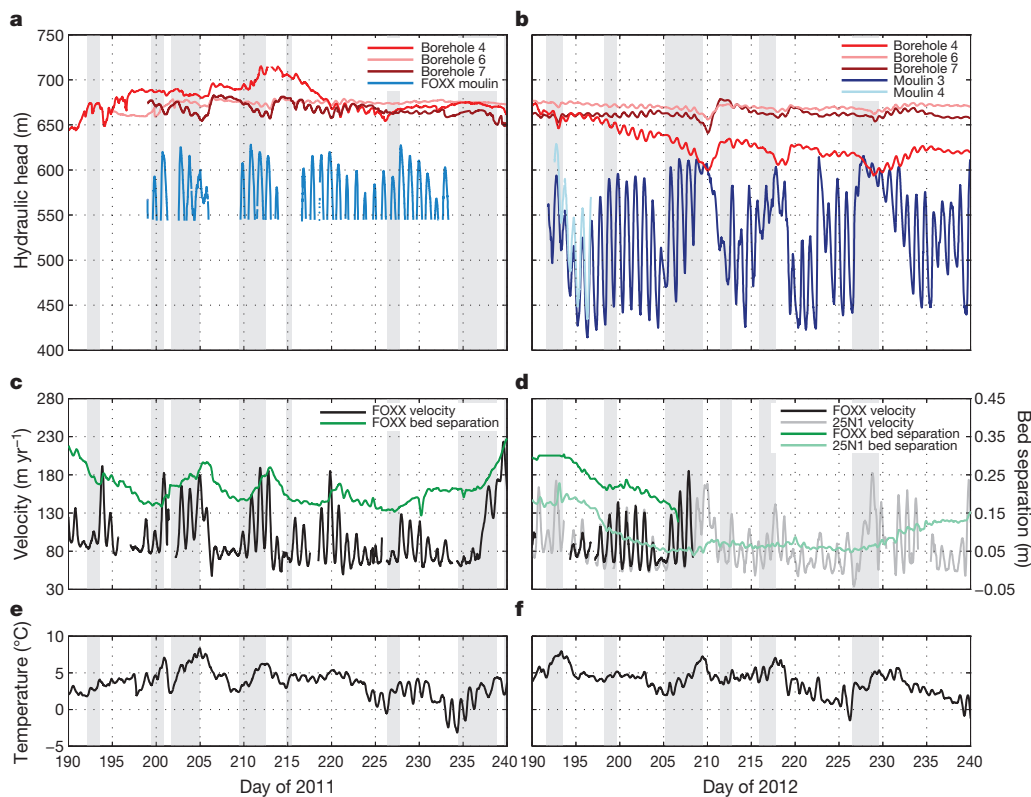
**Figure 1 | Study area in the ablation zone of the western Greenland Ice Sheet.** **a**, Landsat-7 image of Sermeq Avannarlq. Ice-surface contours (marked in metres; brown) are derived from the Greenland Ice Mapping Project (GIMP) surface DEM<sup>29</sup>. Site symbols are indicated in the key. The black box indicates the area in **b**. A 2012 Center for Remote Sensing of Ice Sheets (CREGIS) flight line<sup>30</sup> (green) provides the cross-section in **c**. **b**, 2009 Worldview-2 image of the study area with site locations indicated. **c**, Bed and surface elevations from radar depth sounding<sup>30</sup>. Moulin 3 (navy), FOXX (black) and 25N1 (grey) are projected onto the flight line. Moulin 4 projects onto the FOXX location.

a mean diurnal fluctuation of approximately  $95 \pm 47$  m (about 17% of overburden) during 2012 and minimum values (about 70% of overburden) well below the ice surface. In addition, hydraulic heads in moulins 3 and 4 are synchronous, despite being located in different supraglacial drainage basins and 1.5 km apart (Extended Data Table 2). This similarity in hydraulic heads suggests pressure equalization within

an efficient system<sup>23</sup> that connects these moulins at the bed<sup>3,4</sup>. Further, numerical analysis supports the existence of subglacial channels in our study area (Methods; Extended Data Fig. 3). During periods of steady supraglacial input channel development via meltback may be limited; however, observed melt-event perturbations temporarily increase channel volume, allowing greater transmission of water. From these observations, we infer that moulin hydraulic head reflects subglacial water pressure within a moulin-connected channel system<sup>4,17,18</sup>, which appears to increase in efficiency only over short timescales.

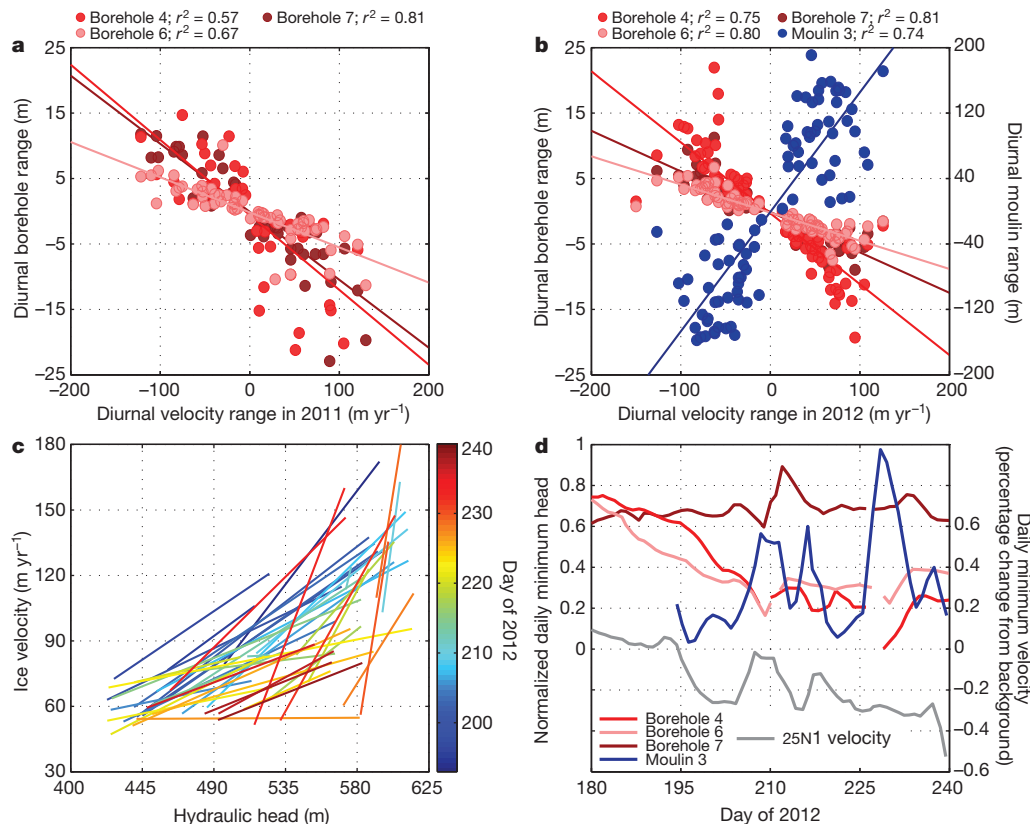
In contrast, our boreholes display high mean hydraulic head (close to or above overburden) and low-amplitude diurnal variability (less than 25 m or <5% of overburden). Borehole hydraulic heads are anti-correlated with ice velocity (Fig. 3a, b; Extended Data Table 2). These systematic differences between moulins and boreholes further suggest that moulins connect to a channelized component of the drainage system, while boreholes monitor an isolated region of the bed unconnected to the channelized system<sup>17,18,21,22</sup> (Methods).

Strong correlations between diurnal peaks in moulin hydraulic heads and ice velocity suggest that pressure variability in the channelized drainage system reduces basal friction in an adjacent active but unchannelized component of the hydraulic system and drives diurnal ice acceleration, as observed in alpine glaciers<sup>14</sup>. However, on longer timescales the relationship between moulin hydraulic head and ice velocity is characterized by hysteresis. In both 2011 and 2012, ice velocity decreases as the melt season progresses, despite relatively constant moulin hydraulic heads (Fig. 3c; Extended Data Fig. 4). Further, neither the minimum nor maximum daily moulin head decreases over the observation period, as would be expected with increasing efficiency, suggesting that pressure decreases in the efficient, channelized system do not control decreases in ice velocity during the latter portion of the melt season (Fig. 3d). Therefore, while variability in moulin head appears to drive diurnal and multi-day velocity



**Figure 2 | Borehole and moulin hydraulic head and ice-surface velocity during 2011 and 2012.** **a**, 2011 hydraulic head measurements from the FOXX moulin (blue), borehole 4 (red), borehole 6 (pink) and borehole 7 (dark red). **b**, 2012 data from moulin 3 (navy) and moulin 4 (light blue). Borehole colours are as in 2011 (**a**). **c**, 2011 GPS-derived ice velocity (black) and bed

separation (green) for FOXX. **d**, 2012 GPS-derived ice velocity for FOXX and 25N1 (grey) and bed separation for FOXX and 25N1 (light green). **e**, **f**, 6-h averaged 2-m air temperature for 2011 and 2012. Grey bars in all panels indicate periods identified as melt events.



**Figure 3 | Relationships between hydraulic head and ice-surface velocity.** **a**, Linear regression between the magnitude of diurnal changes in ice velocity and borehole and moulin hydraulic head for 2011 ( $p < 0.05$ ; borehole 4  $n = 48$ , borehole 6  $n = 59$ , borehole 7  $n = 50$ ). **b**, Linear regression between the magnitude of diurnal changes in ice velocity and borehole and moulin hydraulic head between days 192 and 240 in 2012 ( $p < 0.05$ ; borehole 4  $n = 80$ ,

borehole 6  $n = 79$ , borehole 7  $n = 75$ , moulin 3  $n = 85$ ). Days when borehole hydraulic head and ice velocity are in phase are excluded. **c**, Daily maximum and minimum moulin hydraulic head plotted against associated ice velocity for 2012. **d**, Normalized daily minimum hydraulic heads and ice velocity as a percentage of winter background for 2012. The minimum values are smoothed over 5 days.

variations, longer-term decreases in ice velocity are potentially due to decreasing water pressure in the unchanneled regions of the subglacial hydrologic system.

As observed in alpine glaciers, isolated components of the subglacial hydrologic system may act to resist ice acceleration<sup>21</sup>. Previous studies demonstrate that high (above overburden) water pressures out of phase with ice velocity may be caused by transfer of mechanical support from channelized regions of the bed to isolated regions<sup>7,17,19,22</sup>. However, the diurnal range of borehole hydraulic heads is more strongly anti-correlated with the diurnal range of ice velocity than with the diurnal range of water pressure in nearby active regions of the bed (Methods), suggesting that borehole hydraulic head variability is at least partly the result of non-locally generated sliding<sup>21</sup>. In this proposed mechanism, pressurization of neighbouring regions of the bed that have an efficient connection to the channelized system induces sliding, which is transmitted to these unconnected areas by stress transfer laterally within the ice. In turn, water pressure in unconnected regions of the bed decreases as the volume of bedrock cavities increases through sliding<sup>21</sup> or as basal sediments deform<sup>24,25</sup> without a commensurate water influx. A combination of these processes results in a dynamic water pressure environment, despite the apparent hydraulic isolation of these regions of the bed.

Negative feedback between increased ice velocity and decreased water pressure in unconnected regions can act to limit sliding<sup>20,21</sup> and potentially control minimum ice velocity. This resistance to sliding probably varies both spatially and temporally, owing to changes in connectivity within the isolated system<sup>17,22</sup>. Although borehole hydraulic heads are typically anti-correlated with ice velocity, some boreholes experienced infrequent periods when pressures are in phase with ice velocity following large melt events (for example, on days 211–218 of 2011 in borehole 6,

Fig. 2a). These in-phase periods may indicate ephemeral connections to the unchanneled but interconnected parts of the drainage system, which occur when the hydraulic capacity of the subglacial drainage system is overwhelmed and water flows out of conduits into the surrounding unchanneled system<sup>4,12,14,26</sup>.

While long-term decreases in ice velocity have previously been attributed to decreasing subglacial water pressure caused by increased channelization<sup>1,9,10</sup>, our results suggest that during the latter part of the melt season, the spatial extent of the unchanneled system is a primary control on ice velocity. Basal traction is a function of both the interconnected and isolated regions of the unchanneled system<sup>21</sup>. Therefore, increasing the spatial extent of the interconnected system, at the expense of the isolated system, should result in a larger fraction of the bed at lower water pressures. Gradually increasing the connectivity of the isolated system (that is, opening or enlargement of flow pathways) would have a similar result.

These processes could increase basal traction and decrease ice velocity without a change in the efficiency of the channelized system<sup>21</sup>. Indeed, we observe a gradual decrease in water pressure in two of three boreholes (boreholes 4 and 6; Figs 3d; Extended Data Fig. 1), implying increasing connectivity to active regions of the bed. This process is also observed at a second field site in 2011 (Methods). These reductions in subglacial pressure match well with velocity trends (Fig. 3d), suggesting that systematically decreasing pressures within the isolated system are occurring. The observed decreases in water pressure may result in a higher spatially averaged basal traction at the end of the melt season that persists after meltwater inputs cease. Consequently, this mechanism could explain the winter mediation of summer acceleration<sup>9,10</sup>.

Our results suggest that the subglacial drainage system consists of three components that exert varying control on ice velocity over different



spatiotemporal domains: a moulin-connected channelized system that transports the available meltwater efficiently; an active, interconnected unchannelized system strongly influenced by the channelized system; and an isolated system that responds passively to changes in bed separation because water flow is slow or absent. The spatiotemporal extents of these domains are probably highly variable, with the spatial extent of each component controlled by the spatial distribution of moulin locations<sup>27</sup>, basal and surface topography<sup>28</sup> and the hydraulic conductivity of basal sediments<sup>24,25</sup>.

Direct observations of multiple components of the subglacial hydrologic system illustrate how subglacial drainage efficiency modulates ice dynamics across multiple timescales. The degree of control that each component of the subglacial system exerts on ice velocity depends critically on the period considered. Our data from a sector of the GIS ablation zone indicate that the drainage efficiency of the channelized system does not change substantially during the latter portion of the melt season. Instead, evolving efficiency of non-channelized drainage systems can explain the observed trends in ice velocity during that time. Thus, we caution against the application of channel-only models to explore the seasonal relationship between subglacial hydrology and the ice dynamics of the GIS. Future work should consider the seasonal evolution of all observed components of the subglacial hydrologic system, including isolated regions of the bed. Such investigations are increasing in importance as new results suggest that the melt-season behaviour of the subglacial system affects ice dynamics in the ablation zone<sup>9,10</sup> and, potentially, farther inland<sup>8</sup> after the melt season.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 24 February; accepted 22 August 2014.**

- Bartholomew, I. *et al.* Seasonal evolution of subglacial drainage and acceleration in a Greenland outlet glacier. *Nature Geosci.* **3**, 408–411 (2010).
- Sundal, A. V. *et al.* Melt-induced speed-up of Greenland ice sheet offset by efficient subglacial drainage. *Nature* **469**, 521–524 (2011).
- Chandler, D. M. *et al.* Evolution of the subglacial drainage system beneath the Greenland Ice Sheet revealed by tracers. *Nature Geosci.* **6**, 195–198 (2013).
- Cowton, T. *et al.* Evolution of drainage system morphology at a land-terminating Greenlandic outlet glacier. *J. Geophys. Res.* **118**, 1–13 (2013).
- Lüthi, M., Funk, M., Iken, A., Gogineni, S. & Truffer, M. Mechanisms of fast flow in Jakobshavn Isbræ, West Greenland: Part III. Measurements of ice deformation, temperature and cross-borehole conductivity in boreholes to the bedrock. *J. Glaciol.* **48**, 369–385 (2002).
- Meierbachtol, T., Harper, J. & Humphrey, N. Basal drainage system response to increasing surface melt on the Greenland Ice Sheet. *Science* **341**, 777–779 (2013).
- Ryser, C. *et al.* Sustained high basal motion of the Greenland Ice Sheet revealed by borehole deformation. *J. Glaciol.* **60**, 647–660 (2014).
- Doyle, S. *et al.* Persistent flow acceleration within the interior of the Greenland ice sheet. *Geophys. Res. Lett.* **41**, 899–905 (2014).
- Sole, A. *et al.* Winter motion mediates dynamic response of the Greenland Ice Sheet to warmer summers. *Geophys. Res. Lett.* **40**, 3940–3944 (2013).
- Tedstone, A. J. *et al.* Greenland Ice Sheet motion insensitive to exceptional meltwater forcing. *Proc. Natl Acad. Sci. USA* **110**, 19719–19724 (2013).
- Catania, G. A. & Neumann, T. A. Persistent englacial drainage features in the Greenland Ice Sheet. *Geophys. Res. Lett.* **37**, L02501 (2010).
- Bartholomew, I. *et al.* Short-term variability in Greenland Ice Sheet motion forced by time-varying meltwater drainage: implications for the relationship between subglacial drainage system behavior and ice velocity. *J. Geophys. Res.* **117**, F03002 (2012).
- Hoffman, M. J., Catania, G. A., Neumann, T. A., Andrews, L. C. & Rumrill, J. A. Links between acceleration, melting, and supraglacial lake drainage of the western Greenland Ice Sheet. *J. Geophys. Res.* **116**, F04035 (2011).
- Bartholomew, T. C., Anderson, R. S. & Anderson, S. P. Response of glacier basal motion to transient water storage. *Nature Geosci.* **1**, 33–37 (2008).
- Schoof, C. Ice-sheet acceleration driven by melt supply variability. *Nature* **468**, 803–806 (2010).
- Das, S. B. *et al.* Fracture propagation to the base of the Greenland Ice Sheet during supraglacial lake drainage. *Science* **320**, 778–781 (2008).
- Gordon, S. *et al.* Seasonal reorganization of subglacial drainage inferred from measurements in boreholes. *Hydrol. Processes* **12**, 105–133 (1998).
- Hubbard, B. P., Sharp, M. J., Willis, I. C., Nielsen, M. K. & Smart, C. C. Borehole water-level variations and the structure of the subglacial hydrological system of Haut Glacier d'Arolla, Valais, Switzerland. *J. Glaciol.* **41**, 572–583 (1995).
- Mair, D. *et al.* Hydrological controls on pattern of surface, internal and basal motion during three 'spring events': Haut Glacier d'Arolla, Switzerland. *J. Glaciol.* **49**, 555–567 (2003).
- Hoffman, M. & Price, S. Feedbacks between coupled subglacial hydrology and glacier dynamics. *J. Geophys. Res.* **119**, 411–436 (2014).
- Iken, A. & Truffer, M. The relationship between subglacial water pressure and velocity of Findelengletscher, Switzerland, during its advance and retreat. *J. Glaciol.* **73**, 328–338 (1997).
- Murray, T. & Clarke, G. K. C. Black-box modeling of the subglacial water system. *J. Geophys. Res.* **100**, 10231–10245 (1995).
- Covington, M. D. *et al.* Quantifying the effects of glacier conduit geometry and recharge on proglacial hydrograph form. *J. Hydrol.* **414–415**, 59–71 (2012).
- Dow, C. F. *et al.* Seismic evidence of mechanically weak sediments underlying Russell Glacier, West Greenland. *Ann. Glaciol.* **54**, 135–141 (2013).
- Walter, F., Chaput, J. & Lüthi, M. P. Thick sediments beneath Greenland's ablation zone and their potential role in future ice sheet dynamics. *Geology* **42**, 487–490 (2014).
- Hewitt, I. J. Seasonal change in ice sheet motion due to melt water lubrication. *Earth Planet. Sci. Lett.* **371–372**, 16–25 (2013).
- Gulley, J. D. *et al.* The effect of discrete recharge by moulins and heterogeneity in flow-path efficiency at glacier beds on subglacial hydrology. *J. Glaciol.* **58**, 926–940 (2012).
- Joughin, I. *et al.* Influence of ice-sheet geometry and supraglacial lakes on seasonal ice-flow variability. *Cryosphere* **7**, 1185–1192 (2013).
- Howat, I. M., Negrete, A. & Smith, B. E. The Greenland Ice Mapping Project (GIMP) land classification and surface elevation datasets. *Cryosphere* **8**, 1509–1518 (2014).
- Gogineni, P. *CRISIS Radar Depth Sounder Data* <http://data.crisis.ku.edu/> (2012).

**Acknowledgements** This project was supported by United States National Science Foundation grants OPP-0908156, OPP-0909454 and ANT-0424589 (to CReSIS), Swiss National Science Foundation grant 200021\_127197, and National Geographic Society grant 9067-12. L.C.A. was also supported by UTIG Ewing-Worzel and Gale White Graduate Student Fellowships. M.J.H. was also supported by NASA Cryospheric Sciences and Climate Modeling Programs within the US Department of Energy, Office of Science. J.D.G. was also supported by an NSF Postdoctoral Fellowship (EAR-0946767). Logistical support was provided by CH2MHill Polar Services. The GPS base station and several on-ice GPS units were provided by the UNAVCO facility with support from the NSF and NASA under cooperative agreement EAR-0735156. The University of Minnesota Polar Geospatial Center, funded under NSF OPP collaborative agreement ANT-1043681, provided WorldView imagery. We thank K. M. Schild, J. A. MacGregor, J. D. Nowinski, B. F. Morriss and others for assistance in the field.

**Author Contributions** G.A.C., J.D.G., M.P.L., R.L.H. and T.A.N. designed this study. L.C.A., R.L.H., M.J.H., M.P.L., C.R. and J.D.G. performed the fieldwork. L.C.A. analysed the results and wrote the manuscript. All authors discussed the results and edited the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.C.A. ([landrews@ig.utexas.edu](mailto:landrews@ig.utexas.edu)).

## METHODS

**Seasonal data presentation.** To clearly present diurnal variability in the measured parameters, the time series is shortened to the period of time over which moulin hydraulic heads are measured (Fig. 2). During 2011, velocity and bed separation were measured throughout the melt season; in 2012, all parameters except moulin pressures were recorded over the entire melt season (Extended Data Fig. 1).

**Borehole drilling and instrumentation.** During 2011, 13 boreholes were drilled at two sites, seven at FOXX (Fig. 1; 69° 27' N, 49° 53' W) and six at GULL (69° 27' N, 49° 43' W), using hot water drilling techniques and equipment<sup>5,7,31,32</sup>. Drill water had a consistent temperature and pressure of 80 °C and 8 MPa, respectively. At FOXX, three boreholes were instrumented with pressure transducers at depths between 614 m and 624 m (Extended Data Table 1). In Greenland, thick cold ice results in rapid (less than a day) closure of boreholes, eliminating the influence of surface water input. Further, the volume of the boreholes is assumed to be small relative to the subglacial system. Therefore, we assume that these boreholes function as accurate manometers of the subglacial system.

Two boreholes (boreholes 4 and 6) were instrumented with the Swiss Federal Institute of Technology (ETH) designed digital borehole sensor system (DIBOSS)<sup>7</sup>. Borehole 7 was instrumented with a Geokon 4500HD piezometer. All sensors were connected to the surface via cables able to accommodate an additional 20% strain before breaking (Cortland Cable). Campbell Scientific CR1000 data loggers were used for switching power supply and recording sensor measurements via CFM100 storage modules. The sampling interval ranged from 1 min to 15 min. Data were decimated to 15-min intervals for analysis in this paper. All units remained powered between summer 2011 and spring 2013.

Several observations during and following the drilling process indicate that the boreholes connected to the ice–bed interface: (1) all boreholes reached similar depths of 614 to 624 m; these depths are similar to, though slightly deeper than, depths expected from a 2012 CREStS depth-sounding radar track through the FOXX field site<sup>30</sup>; (2) drilling only ceased when pressures at the drill tip rapidly decreased; (3) boreholes 4 and 6 drained slowly over the course of several hours; and (4) pump tests were performed in boreholes 4 and 6, forcing a connection to the subglacial system. Changes associated with pump tests were ephemeral and boreholes gradually reverted to their pre-pump test state. Borehole 7 also connected to the bed because, even though it did not drain before closure, the diurnal lags and melt event pressure variations in borehole 7 are similar to those observed in boreholes 4 and 6 over the course of both melt seasons.

Borehole hydraulic heads were calculated from measured borehole pressure (from installed pressure transducer), surface elevation and borehole depth:

$$h = \frac{P_w}{\rho_w g} + z_{\text{bed}} \quad (1)$$

where  $h$  is total hydraulic head.  $P_w$  is directly measured from borehole sensors, but is also equal to  $\rho_w g h_w$ . The density of water is  $\rho_w$ ;  $g$  is the acceleration due to gravity and  $h_w$  is water height.  $z_{\text{bed}}$  is the bed elevation determined from GPS-derived ice-surface elevation and measured borehole depth. Representing borehole data as hydraulic head allows us to determine water level absolutely, as measured from sea-level in different locations.

**Moulin instrumentation and measurements.** Moulins were instrumented in both the 2011 (FOXX moulin) and 2012 (moulins 3 and 4) melt seasons (Extended Data Table 1). Water pressures were measured using the Geokon 4500 (2011) or 4500HD (2012) piezometers on armoured cables ranging from 400 m to 600 m in length. Campbell Scientific CR1000 data loggers were used for switching power supply and recording sensor measurements via CFM100 storage modules. Water pressures were corrected for local barometric pressure as measured at the FOXX GPS station (2011) and moulin 3 (2012). Sampling intervals were 5 min or 15 min. Data were decimated to 15 min for analysis except where noted.

Moulin instrumentation is complicated by moulin geometry and water-level fluctuations that occur during pressure sensor installation. To constrain the absolute elevation of the pressure sensor in each moulin several corrections to the measured moulin water level must be made. These corrections include adjustment for the rise of water levels during the sensor installation and the mean slope of the conduit.

During sensor installation, we periodically paused while lowering to observe the water-level rise with the sensor held in a static position. These pauses allowed us to constrain the total change in water level over the course of sensor installation. Once the water level was corrected to a static level, we corrected for the slope of the moulin. However, it is important to note that moulins in Greenland are generally nearly vertical<sup>11</sup>, so this correction is small. We subtracted the corrected sensor depth from the GPS-derived surface elevation at each site to constrain the vertical sensor location. Sensor elevation and measured water level provide hydraulic head without the need to use poorly constrained ice thicknesses and bed elevations at moulin sites.

Owing to the possibly complex geometry of moulins, the error in sensor location is higher than the error in borehole-sensor locations. We estimate the error in absolute

head measurements to be approximately 20 m. Uncertainty related to absolute head measurements does not affect measurements of relative changes, such as diurnal amplitudes of hydraulic head. The sensor in moulin 3 is near the bed of the ice sheet, as indicated by the continual increase in water as the sensor was lowered and because the total cable length deployed was similar to the local ice thickness. The sensor in moulin 4 is approximately 175 m from the bed, on the basis of similar observations. The 2011 sensor in the FOXX moulin is much shallower, as indicated by the truncated pressure record. Owing to the deployment techniques, the absolute sensor location is not as well constrained in 2011 as in 2012. We estimate the FOXX sensor location to be 455 m above the bed.

Moulins connect directly to highly efficient components of the subglacial system<sup>27</sup> and have previously been used with varying levels of success to measure the water level of the channelized component of the subglacial system<sup>4,33–35</sup>. Although moulins are not considered manometers, measured moulin pressures can be considered equivalent to subglacial water pressures within the channelized system, because pressure changes in multiple moulins are coincident (Extended Data Table 2), despite drainage basins of differing sizes and discharges. The volume of water in a moulin is large relative to the volume of water being discharged at the bed, so water flow within the moulin is slow<sup>36</sup>; we therefore neglect the velocity head.

In addition, to be considered effective measures of water pressure, the volume of water within the moulin must be small when compared to the volume of water in the channelized system. Considering that large portions of the channelized system are connected<sup>3,4</sup>, the total volume in a single moulin is probably small relative to the total water volume within the channelized system. With these assumptions, we consider the measured moulin pressures to be representative of subglacial pressures in the efficient component of the subglacial system.

**Coincident observations.** Over the course of 4 days (approximately day 192.5 to day 196.5), we were able to monitor the water level in two moulins concurrently (Fig. 2b). These observations suggest that the hydraulic heads of both moulins behave very similarly, with peaks, and even small perturbations, occurring in both moulins at the same time (Extended Data Table 2). We did not calculate the hydraulic gradient between the two moulins because the length of the subglacial channel is unknown and channel paths may diverge from those predicted by hydraulic potential theory<sup>27</sup>. However, we do use the head difference as a proxy for the hydraulic gradient<sup>17</sup>, assuming that the subglacial channel does not alter its path significantly over the course of the melt season.

**Ice velocity and bed separation.** Kinematic GPS positions were determined using carrier-phase differential processing relative to a bedrock mounted reference station using Track 1.24<sup>37</sup> and techniques described by ref. 13. GPS observations at all stations were logged at 15-s intervals, and the relative position of each on-ice station was determined at this frequency. Each 15-s time series was smoothed with a 6-hour moving average, applied to reduce spurious signals associated with GPS uncertainties, and then decimated to 15-min time series. Using the mean error generated during processing, the horizontal and vertical position errors for 2011 (and 2012) were  $\pm 9.9$  cm ( $\pm 8.8$  cm) and  $\pm 1$  cm ( $\pm 1$  cm) respectively, with a velocity uncertainty of  $\sim 8.8$  m yr<sup>−1</sup> ( $\sim 7.7$  m yr<sup>−1</sup>).

During 2012, the antenna pole for the FOXX GPS station melted out owing to higher than expected ablation rates. The data become unreliable after day 208, and the station ceased recording by day 215. Owing to the similarity between FOXX and 25N1 data (Fig. 2; Extended Data Fig. 1), the 25N1 GPS data were used for analysis of 2012 data. Over the entire melt season, ice velocities at both stations display the characteristic decrease in daily minima<sup>13</sup> (Extended Data Fig. 1).

Bed-separation calculations were performed incorporating additional observations from nearby GPS stations and following the procedures described in detail in previous work<sup>13</sup>. Both longitudinal (along flow) and lateral (across flow) strain rates  $\dot{\epsilon}$  were calculated from GPS data as follows:

$$\dot{\epsilon} = \frac{1}{l_0} \frac{\Delta l}{\Delta t} \quad (2)$$

where  $\Delta l$  is the change in baseline distance between stations,  $\Delta t$  is the change in time between measurements and  $l_0$  is the initial baseline distance. Longitudinal strain at our field location is generally compressional over the course of the season. During 2012, GPS malfunction prevented the collection of data from 19N1 and decreased the availability of data from the FOXX GPS; bed separation was therefore calculated between 25N1 and GULL and is presented as a proxy (Fig. 2 and Extended Data Fig. 1). During 2012, malfunction of a GPS station used to determine lateral strain between 25N1 and GULL required that we assume a constant lateral strain for bed separation calculations after day 205.

We approximate the vertical strain rate  $\dot{\epsilon}_z$  with the continuity equation, assuming ice is incompressible:

$$\dot{\epsilon}_{zz} = -(\dot{\epsilon}_{xx} + \dot{\epsilon}_{yy}) \quad (3)$$

where  $x$ ,  $y$  and  $z$  represent the longitudinal, lateral and vertical directions, respectively.

Following accepted methodology<sup>38–43</sup>, we decompose the vertical motion measured by the GPS as:

$$w_s = u_b \tan(\alpha) + \dot{\epsilon}_{zz} H + \dot{c} \quad (4)$$

where  $u$  is the horizontal velocity,  $\alpha$  is the bed slope,  $H$  is the ice thickness at the GPS station, as measured by borehole depth, and  $\dot{c}$  is the rate of bed separation (cavity opening and/or till dilation). Subscripts 's' and 'b' refer to the surface and bed of the ice sheet, respectively. Equation (4) is solved for  $\dot{c}$  using observations of the other parameters.

The length scale over which bed slope should be measured is difficult to estimate owing to the variability of bed roughness over several ice thicknesses in our study region<sup>12,44</sup>. Therefore, we chose to derive the bed slope from calculations before the onset of summer melting and the associated increase in velocity. During this window, we assume that vertical strain and the rate of cavity opening are constant, yielding:

$$\alpha = \tan^{-1} \left( \frac{w_{s,bg} - \dot{\epsilon}_{zz,bg} H}{u_{b,bg}} \right) \quad (5)$$

where the subscript 'bg' represents background conditions.

We note that diurnal variations in bed separation are generally within the range of error. Our results suggest that vertical strain is a non-negligible contribution to vertical motion in regions of the GIS (Extended Data Fig. 2); simply removing the bed-parallel component of bed separation while assuming that  $\dot{\epsilon}_{zz}$  is negligible may result in inaccurate estimates of bed separation.

**Melt events.** During 2011 and 2012, an automatic weather station measured a wide range of atmospheric conditions every 5 min, including ablation, incoming and reflected short-wave radiation, wind speed and direction, humidity and the air temperature at a height of 2 m above the ice surface. To quantitatively constrain melt events, we simply difference consecutive calculations of 24-hour average temperature (with noon Coordinated Universal Time (UTC) as the midpoint). A positive temperature differential (that is, an increase in the 24-hour average temperature) of 0.5 °C is considered a melt event. To limit visual confusion, the start of the melt event is plotted as the minimum of the first day and the maximum of the last day of the melt event (Fig. 2).

**Cross-correlation analysis.** To characterize the lead–lag relationship between borehole hydraulic head, moulin hydraulic head and ice velocity, we performed cross-correlation analyses. We used the maximum cross-correlation coefficients and associated lags to examine the temporal relationship between various time series<sup>17</sup> (Extended Data Table 2).

The ice velocities used for cross-correlation analysis have a 1-hour moving average applied to reduce spurious noise<sup>13</sup> but still maintain independence between velocity measurements used in this analysis. We then decimated ice velocity and hydraulic head measurements to a 1-hour sampling interval. For moulins 3 and 4, we use a 5-min sampling interval to determine more closely a potential lagged relationship. We detrended all data using a 24-hour moving window mean and recorded measurements as standardized residuals<sup>17</sup>. Data gaps, which primarily occur in velocity data (~3% during both years), are filled using linear interpolation. Less than 1% (and generally 0%) of borehole and moulin data were linearly interpolated in either year; the exception being the FOXX moulin data, with 6% in 2011. All re-expressed time series, except the FOXX moulin, approximate a Gaussian distribution with a mean centred at zero. Because the FOXX moulin does not approximate a Gaussian distribution, we do not present cross-correlation analysis that includes the FOXX moulin.

Borehole hydraulic head and ice velocity exhibit the strongest inverse correlation with a lag of ~4 hours. As 1-hour averaged ice velocities have inherently higher errors than 6-hour averaged ice velocities, correlation coefficients with ice velocity are low but still significant (Extended Data Table 2). Moulins 3 and 4 are strongly correlated at zero lag (Extended Data Table 2), despite having differing supraglacial inputs and geometries, suggesting pressure equalization within the channelized subglacial system<sup>23</sup>.

**Subglacial channel geometry.** Recent evidence suggests that characteristics of the GIS ablation zone distal from the margin (low surface slope and limited conduit meltback) prevent the formation of an extensive channelized system<sup>6</sup>. However, dye tracing suggests the presence of channelization through much of the ablation area<sup>3</sup>. Our moulin observations indicate a component of the subglacial system that is efficiently conducting the available melt water. However, this may be the result of an efficient distributed pathway<sup>6</sup>. To characterize the nature of the efficient system we performed a simple numerical analysis to explore the channel stability in our study area.

Using the hydraulic head of moulin 3 and supraglacial discharge estimated from ablation measurements at the FOXX weather station, we calculated the change in channel geometry over ~30 days in 2012, following previous work<sup>6,15</sup>:

$$\frac{dS}{dt} = c_1 Q \Psi + u_b h_b - c_2 N^n S \quad (6)$$

where subglacial discharge  $Q$  is calculated as a function of moulin head  $h_m$  and supraglacial input,  $Q_{sur}$ :

$$Q = Q_{sur} - \frac{dh_m}{dt} S_m \quad (7)$$

and  $\Psi$  is the hydraulic gradient calculated as a function of downstream (indicated by subscript 'j') hydraulic head:

$$h_j = h_m + z_j - z_m - \frac{l_{m,j} c_3}{\rho_w g} Q^2 S_{m,j}^{-5/2} \quad (8)$$

Effective pressure  $N$  was calculated at the point midway between  $h_m$  and  $h_j$ . Additional parameters are listed in Extended Data Table 3.

For these calculations, we set  $u_b h_b = 0$  to clarify the role of creep closure and channel melt-back in the maintenance of subglacial channels. This assumption is applicable for our data set because we measure pressures after the assumed onset of channelization<sup>4</sup>. Once channels are larger than bedrock bumps, cavity opening due to sliding probably plays a very limited part in continued channelization<sup>15</sup>. Initial subglacial channel size is determined by initiating the model with a small channel and sinusoidal inputs approximating the daily average and range of moulin 3's head and supraglacial discharge into the modelled moulin until the channel size stabilizes. With this approach we cannot specifically address the timescale of channel development because we do not constrain supraglacial input and moulin hydraulic head at the beginning of the melt season. After determining the initial channel size, we forced the system with observations of supraglacial input and the hydraulic head of moulin 3 (Extended Data Fig. 4). Supraglacial input is calculated by scaling ablation measurements to half of the moulin 3 drainage basin area<sup>45</sup>. As moulin hydraulic head exceeds floatation only rarely, we assumed that all water entering the moulin enters and stays within the channelized system over the short distance assessed here. This assumption is conservative with respect to maximum channel melt back.

Using a distance (1.5 km) and a bed and surface slope similar to those of our study area, we were able to qualitatively reproduce the head difference between moulins 3 and 4 using this simple numerical analysis (Extended Data Fig. 4). Although this analysis cannot directly address the timescale of channel development, it does suggest that with the observed supraglacial input and moulin hydraulic head, subglacial channels can be maintained in our study area. However, we do note that supraglacial melt perturbations may be required to sustain channels over long periods; so steady-state modelling may be insufficient to explore the spatial extent of subglacial channels under the GIS. These calculations suggest that periods with elevated hydraulic gradients, probably during and shortly after melt events, allow for increased melting to effectively counter creep closure (Extended Data Fig. 4). This melting results in channel enlargement; however, during periods of normal diurnal variability, the available supraglacial melt and the associated hydraulic gradient are generally lower, resulting in rapid channel closure. Regular melt events may be essential in the maintenance of subglacial channels in this region of the GIS.

Recent studies indicate that increased efficiency of subglacial channels during high melt years could result in the observed decline in late season ice velocities through more extensive drainage of the bed following the cessation of surface melting<sup>9,10</sup>. However, our results suggest that the rapid adjustment of subglacial channels to the available melt, and the need for melt events to open the conduits periodically, may preclude this possibility. Instead, the observed declines in ice velocity<sup>1,9,10,13</sup>, may be the result of changes in unchannelized regions of the subglacial system.

**Potential mechanisms for borehole diurnal variability.** Although the boreholes in this study connected to the bed, there are several lines of evidence that suggest borehole pressures are responding to pressure variations due to basal sliding rather than the direct propagation of water pressure pulses at the bed. First, during 2011 (and mostly during 2012), borehole hydraulic heads are always higher than those of the moulin that is only 300 m away, thus preventing the propagation of water from the channelized system to the borehole locations. Second, if pressure fluctuations in our boreholes resulted from propagation of diurnal pulses from channels to boreholes through till, we would expect boreholes further from the channel to exhibit smaller and more lagged peak diurnal water pressures. In contrast, we find that water-pressure maxima and minima in boreholes display no consistent pattern in lag times (Fig. 2), and cross-correlation analysis indicates that lag does not increase as a function of distance from the moulin in 2011 (Extended Data Table 2). Third, we computed the hydraulic diffusivity needed to produce the observed 15-h phase lag between moulin and borehole pressure peaks following ref. 18:

$$D = \frac{x^2}{2\omega t^2} \quad (9)$$

where  $x$  and  $t$  are the distance and time lag, respectively, between the channel and the borehole sensor and  $\omega$  is the angular frequency of the periodic boundary condition ( $\omega = 7.27 \times 10^{-5} \text{ s}^{-1}$  for a diurnal cycle). This calculation results in a diffusivity of approximately  $10^{-1} \text{ m}^2 \text{ s}^{-1}$ , several orders of magnitude larger than



expected hydraulic diffusivities for glacial till<sup>46</sup> and in the range of the hydraulic diffusivity for rock<sup>47</sup> and yet there is evidence of thick sediments beneath our field area<sup>25</sup>. Fourth, borehole minima are more closely correlated to moulin and velocity maxima (Extended Data Table 2). As a result, we do not believe that the pressure lag observed in the borehole record results from diffusion of the moulin-generated pressure wave through subglacial sediments. These observations and recent modelling results<sup>7</sup> lead us to conclude that borehole-pressure fluctuations result from non-locally forced ice motion. In this case, we expect and observe the pressure fluctuations in the boreholes to scale with diurnal velocity variability (Fig. 3).

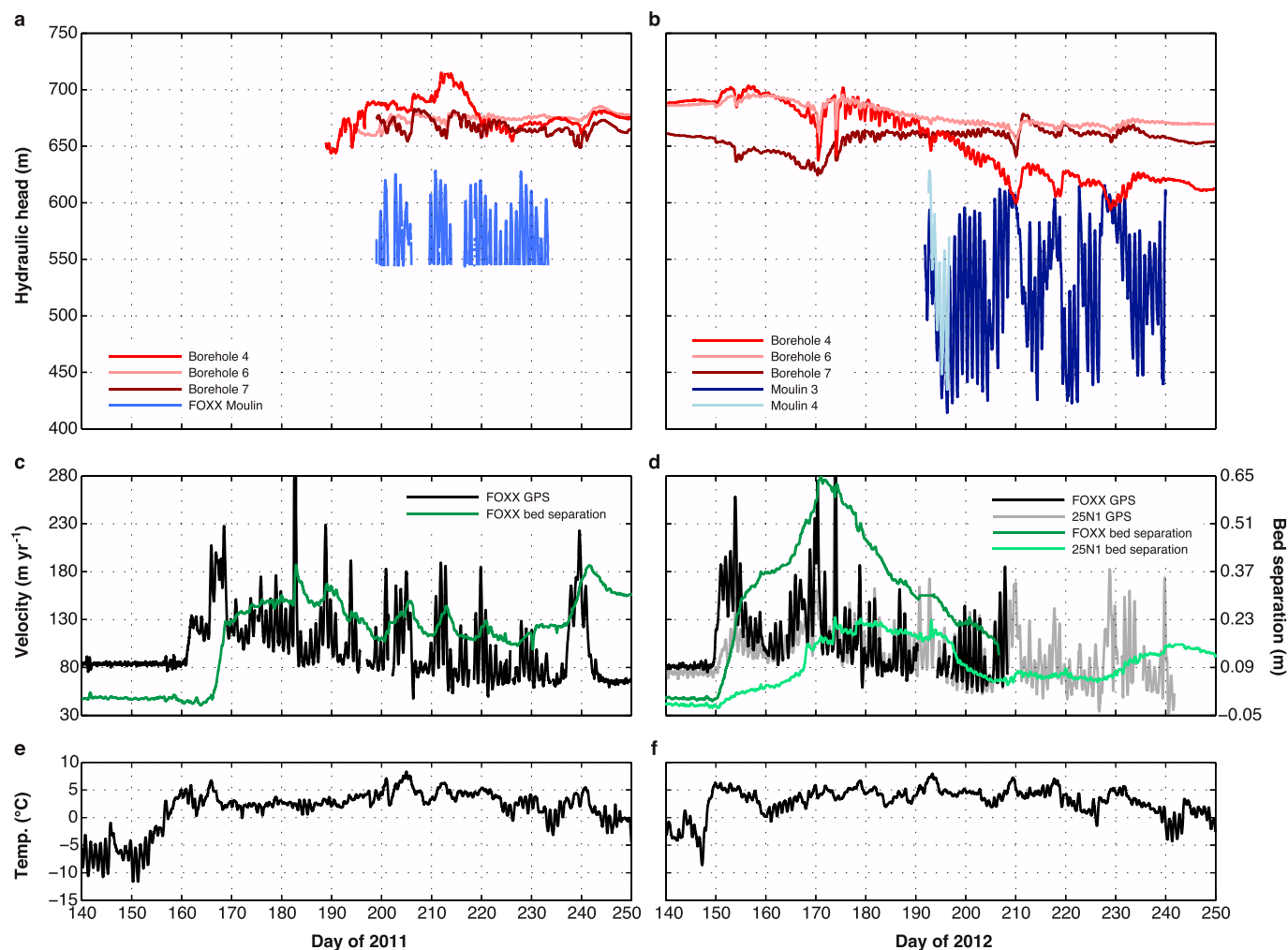
**Flow coupling and mechanical support transfer.** To constrain the mechanism driving changing pressures measured via boreholes, we examine the relationship between the diurnal range in borehole hydraulic head and the diurnal ranges in both ice velocity and moulin hydraulic head. Borehole water levels are anti-correlated with horizontal ice velocity and moulin water level (Figs 2 and 3a, b). Such anti-correlated behaviour has been observed on alpine glaciers and is generally hypothesized to occur through the transfer of mechanical support between the efficient and weakly connected regions of the bed<sup>17–19,22</sup>. In our study, the two components are represented by the moulin and borehole water levels, respectively. However, if load transfer is the only mechanism controlling borehole water pressure, diurnal changes in borehole water level should be most strongly anti-correlated with diurnal changes in moulin water level. Instead, the diurnal magnitude of borehole head is more strongly anti-correlated with the diurnal magnitude of ice velocity, as follows. For moulin 3, (~day 192–240), borehole 4  $r^2 = 0.68$  ( $p < 0.05$ ;  $n = 89$ ), borehole 6  $r^2 = 0.70$  ( $p < 0.05$ ;  $n = 88$ ), borehole 7  $r^2 = 0.72$  ( $p < 0.05$ ;  $n = 85$ ). For ice velocity, (~day 192–240), borehole 4  $r^2 = 0.75$  ( $p < 0.05$ ;  $n = 80$ ), borehole 6  $r^2 = 0.80$  ( $p < 0.05$ ;  $n = 79$ ), borehole 7  $r^2 = 0.81$  ( $p < 0.05$ ;  $n = 75$ ). Local load transfer would also result in moulin water levels and borehole pressures to be directly out of phase<sup>22</sup>. We observe that minimum daily borehole pressures consistently lag daily maximum ice velocity and moulin water level by about 4 hours (Extended Data Table 2), further suggesting that flow coupling may be important in controlling unchanneled regions of the bed<sup>21</sup>.

**Moulin-velocity hysteresis.** To examine the relationship between moulin hydraulic head and ice velocity over the course of the observation period, we plot moulin hydraulic head against ice velocity. During 2011 and 2012, the relationship between moulin hydraulic head and ice velocity changes, showing a decreasing trend in velocity without a similar change in moulin hydraulic head. Diurnal hysteresis is evident in both years (Extended Data Fig. 4).

**GULL borehole observations.** During the 2011 field season, a series of boreholes were also drilled and instrumented at GULL (69° 27' N, 49° 43' W), a site approximately 5 km up the flow line from FOXX. These boreholes were less than 500 m downstream of a moulin slowly draining a supraglacial lake. As at FOXX, pressure transducers were installed in three boreholes; however, owing to a highly deformable layer of ice, sensor cables sheared before the start of the 2012 melt season<sup>7</sup>.

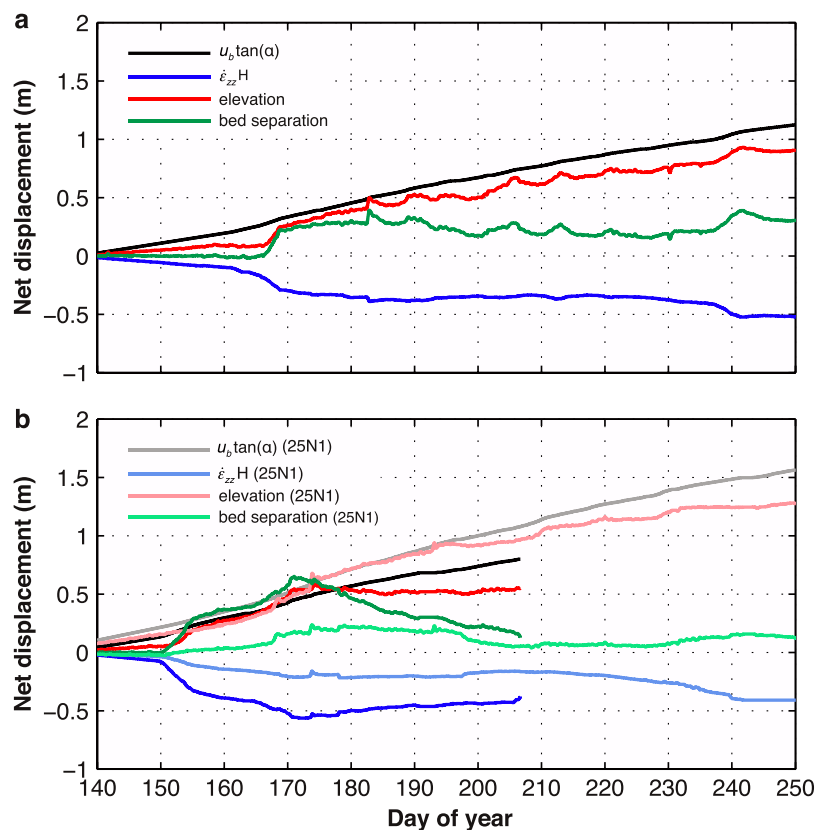
Borehole head at GULL demonstrates small diurnal variations that are out of phase with velocity. Yet, despite a temporally limited water pressure record, we observe similar trends in decreasing daily minimum velocity and borehole head values and borehole head as observed at FOXX during the 2012 melt season (Extended Data Fig. 5). Although these observations are limited by the borehole spatial distribution, this trend suggests that these long-term dynamic changes may be widespread.

31. Iken, A., Echelmeyer, K. & Harrison, W. D. In *Ice Core Drilling* (eds Rado, C. & Beaudoin, D.) *Proc. Third Int. Workshop on Ice Drilling Technology (Grenoble, 10–14 October)* 123–134 (Laboratoire de Glaciologie et Géophysique de l'Environnement, 1988).
32. Humphrey, N. & Echelmeyer, K. Hot-water drilling and bore-hole closure in cold ice. *J. Glaciol.* **36**, 287–298 (1990).
33. Holmlund, P. & Hooke, R. LeB. High water-pressure events in moulins, Storglaciären, Sweden. *Geogr. Ann.* **A 65**, 19–25 (1983).
34. Iken, A. Measurements of water pressure in moulins as part of a movement study of the White Glacier, Axel Heiberg Island, Northwest Territories, Canada. *J. Glaciol.* **11**, 53–58 (1972).
35. Vieli, A., Jania, J., Blatter, H. & Funk, M. Short-term velocity variations on Hansbreen, a tidewater glacier in Spitsbergen. *J. Glaciol.* **50**, 389–398 (2004).
36. Werder, M. A., Loye, A. & Funk, M. Dye tracing a jökulhlaup: I. Subglacial water transit speed and water-storage mechanism. *J. Glaciol.* **55**, 889–898 (2009).
37. Chen, G. *GPS Kinematic Positioning for Airborne Laser Altimetry at Long Valley, California*. PhD thesis, MIT (1998); <http://dspace.mit.edu/handle/1721.1/9680>.
38. Mair, D., Nienow, P., Sharp, M., Wohlleben, T. & Willis, I. Influence of subglacial drainage system evolution on glacier surface motion: Haut Glacier d'Arolla, Switzerland. *J. Geophys. Res.* **107**, <http://dx.doi.org/10.1029/2001JB000514> (2002).
39. Mair, D., Sharp, M. J. & Willis, I. C. Evidence for basal cavity opening from analysis of surface uplift during a high-velocity event: Haut Glacier d'Arolla, Switzerland. *J. Glaciol.* **48**, 208–216 (2002).
40. Anderson, R. L. *et al.* Strong feedbacks between hydrology and sliding of a small alpine glacier. *J. Geophys. Res.* **109**, F03005 (2004).
41. Sugiyama, S. & Gudmundsson, H. Short-term variations in glacier flow controlled by subglacial water pressure at Lauteraargletscher, Bernese Alps, Switzerland. *J. Glaciol.* **50**, 353–362 (2004).
42. Harper, J. T., Humphrey, N. F., Pfeffer, W. T. & Lazar, B. Two modes of accelerated glacier sliding related to water. *Geophys. Res. Lett.* **34**, L12503 (2007).
43. Howat, I. M., Tulaczyk, S., Waddington, E. & Björnsson, H. Dynamic controls on glacier basal motion inferred from surface ice motion. *J. Geophys. Res.* **113**, F03015 (2008).
44. Catania, G. A., Neumann, T. A. & Price, S. F. Characterizing englacial drainage in the ablation zone of the Greenland Ice Sheet. *J. Glaciol.* **54**, 567–578 (2008).
45. McGrath, D., Colgan, W., Steffen, K., Lauffenburger, P. & Balog, J. Assessing the summer water budget of a moulin basin in the Sermeq Avannarleq ablation region, Greenland ice sheet. *J. Glaciol.* **57**, 954–964 (2011).
46. Paterson, W. S. B. *The Physics of Glaciers* 3rd edn, 60 (Pergamon, 1994).
47. Li, V. Estimation of in-situ hydraulic diffusivity of rock masses. *Pure Appl. Geophys.* **122**, 545–559 (1985).



**Extended Data Figure 1 | Borehole and moulin head and ice-surface velocity over two melt seasons.** **a**, 2011 measurements from FOXX moulin (blue), borehole 7 (dark red), borehole 6 (pink) and borehole 4 (red). **b**, 2012 measurements from moulin 3 (navy) and moulin 4 (light blue). **c**, 2011 ice velocity (black) and bed separation (green) for FOXX. Peak velocity on day 182

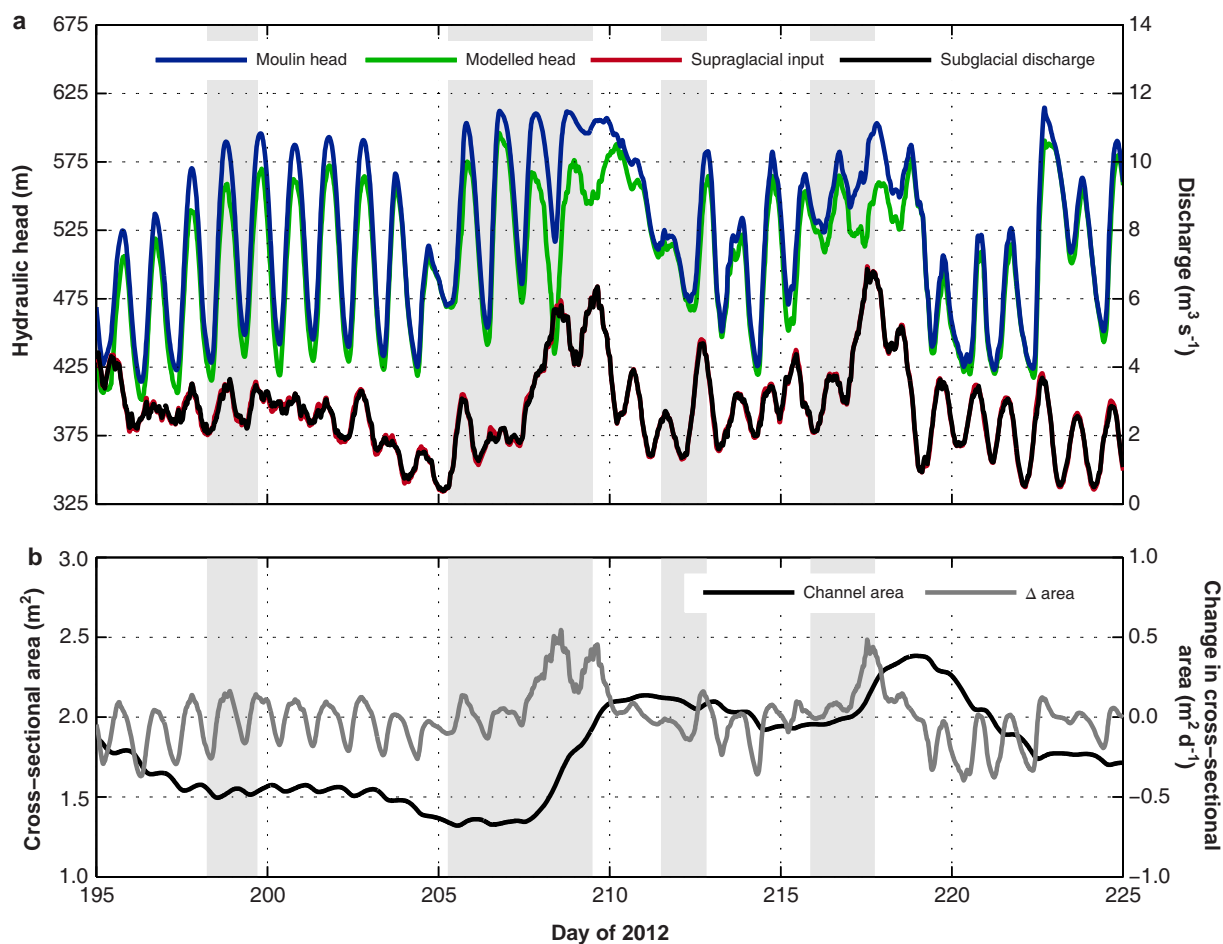
is 402.3 m yr<sup>-1</sup> (exceeding the y-axis limit). **d**, 2012 ice velocity and bed separation for FOXX and 25N1 (grey, light green). Peak velocity for FOXX and 25N1 (312 m yr<sup>-1</sup> and 337 m yr<sup>-1</sup>) occurred on day 173. **e**, **f**, 6-h averaged air temperature for 2011 and 2012. Grey bars are melt events.



**Extended Data Figure 2 | Components of vertical motion 2011 and 2012.**  
**a**, Components of vertical motion for 2011 at FOXX. Bed parallel motion ( $u_b \tan(\alpha)$ ; black), strain thickening and thinning ( $\epsilon_{zz} H$ ; blue), elevation with

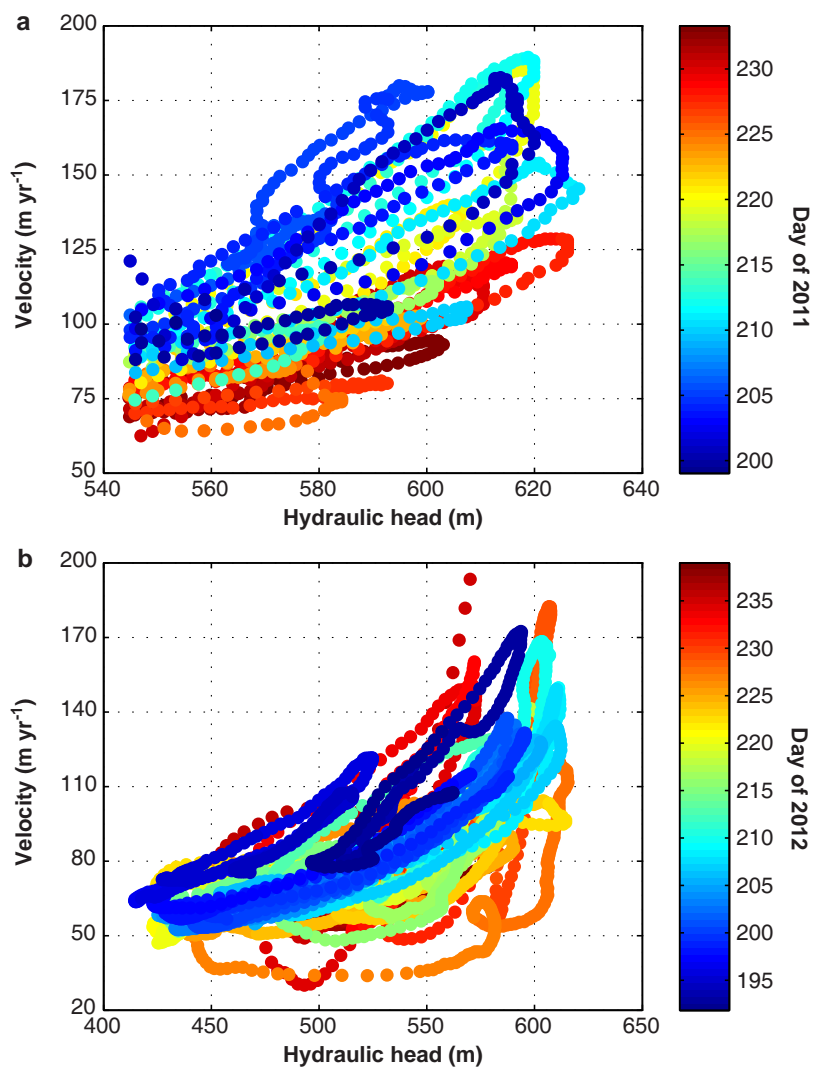
measured winter elevation removed (red), and calculated bed separation (green). **b**, Components of vertical motion for 2012. Colours as in **a** for FOXX. Lighter colours correspond to components of vertical motion from 25N1.





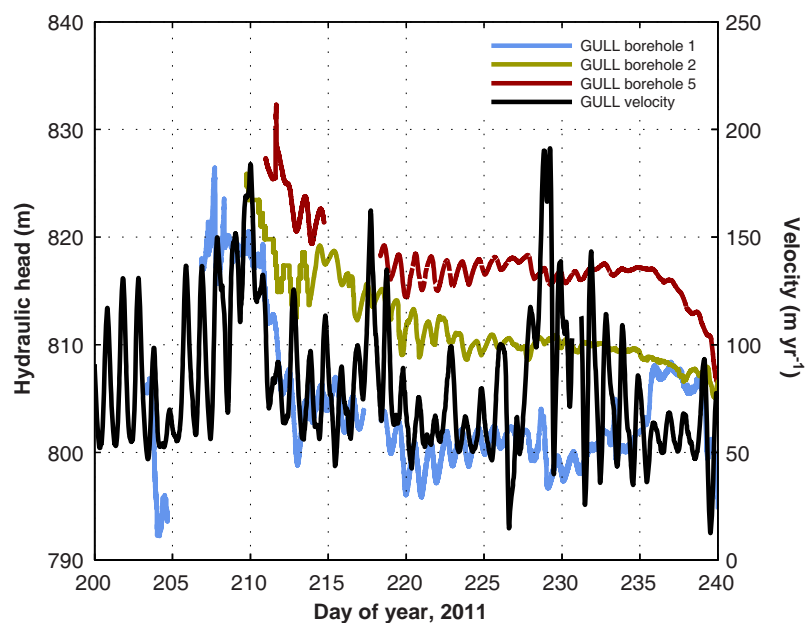
**Extended Data Figure 3 | Modelled hydraulic head and conduit geometry.**  
**a,** Moulin 3 head (blue) and supraglacial input (red) are model inputs. Predicted downstream head is calculated from equation (8) (green). Subglacial discharge (black) is calculated as a function of head change and supraglacial

inputs. It does not vary significantly from supraglacial input even when a large ( $\sim 20 \text{ m}^2$ ) moulin geometry is used. **b,** Modelled subglacial channel cross-sectional area (black) changes rapidly (grey) during and shortly after expected melt events (grey bars).



**Extended Data Figure 4 | Seasonal relationship between moulin head and ice velocity for 2012 and 2011.** Moulin hydraulic head and associated ice velocity data plotted every 15 min over the course of the measurement periods

for 2011 (a) and 2012 (b). 2011 data are truncated below 543 m by the high elevation of the moulin sensor.



**Extended Data Figure 5 | Borehole hydraulic heads and ice velocity at GULL during 2011.** Three hydraulic head records (red, yellow, blue) from boreholes located  $\sim 0.5$  km from a moulin. Ice velocity from GULL GPS (black),  $\sim 0.75$  km south of GULL boreholes.



**Extended Data Table 1 | Site coordinates and characteristics**

Field site	Latitude	Longitude	Elevation (m)	Ice thickness (m)	Surface characteristics
FOXX GPS & weather station	69.4458	-49.8847	706	602*	uncrevassed, in small basin
25N1 GPS	69.4454	-49.7890	851	608*	local crevasses
Borehole 4	69.4456	-49.8802	706	624	uncrevassed, near large supraglacial stream
Borehole 6	69.4463	-49.8807	706	614	uncrevassed, near large supraglacial stream
Borehole 7	69.4464	-49.8805	706	623	uncrevassed, near large supraglacial stream
FOXX moulin	69.4446	-49.8859	703	620*	supraglacial stream entering from the northeast
Moulin 3	69.4358	-49.9092	657	564*	supraglacial streams entering from the east & west
Moulin 4	69.4296	-49.8785	709	540*	supraglacial stream entering from the east

Location and surface elevations determined from GPS. Ice thicknesses at borehole locations determined during drilling. \*Ice thicknesses interpolated from CReSIS radar sounding data<sup>30</sup>.

Extended Data Table 2 | Cross-correlation analysis

	2011		2012			
	Observation period (day 195 to 235)		Melt season (day 150 to 240)		Moulin period (day 192 to 240)	
	Peak to minimum	Peak to peak	Peak to minimum	Peak to peak	Peak to minimum	Peak to peak
Borehole 4/ ice velocity	-0.15 (7)	0.16 (18)	-0.45 (5)	0.38 (16)	-0.49 (4)	0.46 (16)
Borehole 6/ ice velocity	-0.21 (1)	0.23 (17)	-0.38 (5)	0.34 (16)	-0.49 (3)	0.46 (15)
Borehole 7/ ice velocity	-0.39 (4)	0.39 (15)	-0.42 (5)	0.39 (16)	-0.44 (4)	0.43 (16)
Moulin 3/ ice velocity	--	--	--	--	0.48 (0)	-0.45 (-13)
Borehole 4/ moulin 3	--	--	--	--	-0.83 (4)	0.75 (16)
Borehole 6/ moulin 3	--	--	--	--	-0.80 (2)	0.74 (14)
Borehole 7/ moulin 3	--	--	--	--	-0.76 (3)	0.69 (16)
Moulin 4/ moulin 3	--	--	--	--	1.00 (0)	-0.80 (11.5)

Maximum positive and negative correlation coefficient between indicated data sets; lag times are noted in parentheses. In 2011, sample sizes for borehole 4, borehole 6 and borehole 7 cross-correlations are  $n = 1,105, 981$  and  $839$ , respectively. For the 2012 melt season,  $n = 2,159$ . Between days 192 and 240 of 2012,  $n = 1,155$ . The moulin 4 to moulin 3 cross-correlation sample size is  $n = 901$ . The 99% confidence interval for all cross-correlation coefficients is less than  $0 \pm 0.1$ . Negative lags indicate that the first series leads the second series.

Extended Data Table 3 | Parameters used in conduit-geometry calculations

Symbol	Value	Parameter
A	$5.3\text{e-}24 \text{ Pa}^{-3} \text{ s}^{-1}$	Glen's flow law coefficient
$c_t$	$7.5\text{e-}8 \text{ J kg}^{-1} \text{ K}^{-1}$	Pressure melting coefficient
$c_w$	$4.22\text{e}3 \text{ J kg}^{-1} \text{ K}^{-1}$	Specific heat capacity, ice
$f$	0.1	Friction factor
n	3	Glen's flow law exponent
L	$3.35\text{e}5 \text{ J kg}^{-1}$	Latent heat of fusion
$\rho_i$	$910 \text{ kg m}^{-3}$	Density of ice
$\rho_w$	$1,000 \text{ kg m}^{-3}$	Density of water
$h_b$	m	Bedrock bump height
S	$\text{m}^2$	Conduit cross-sectional area
$S_m$	$\text{m}^2$	Moulin cross-sectional area
$u_b$	$\text{m a}^{-1}$	Basal sliding velocity
z	m	Bed elevation
$\Psi$	$\text{Pa m}^{-1}$	Hydraulic gradient
N	$\rho_i g h_i - \rho_w g h_w$	Effective pressure
$c_1$	$(1 - \rho_w c_t c_w) / \rho_i L$	Melting parameter
$c_2$	$2An^{-n}$	Closing parameter
$c_3$	$(\pi + 2)\rho_w f / 2^{5/2} \pi^{1/2}$	



# Prevalence of viscoelastic relaxation after the 2011 Tohoku-oki earthquake

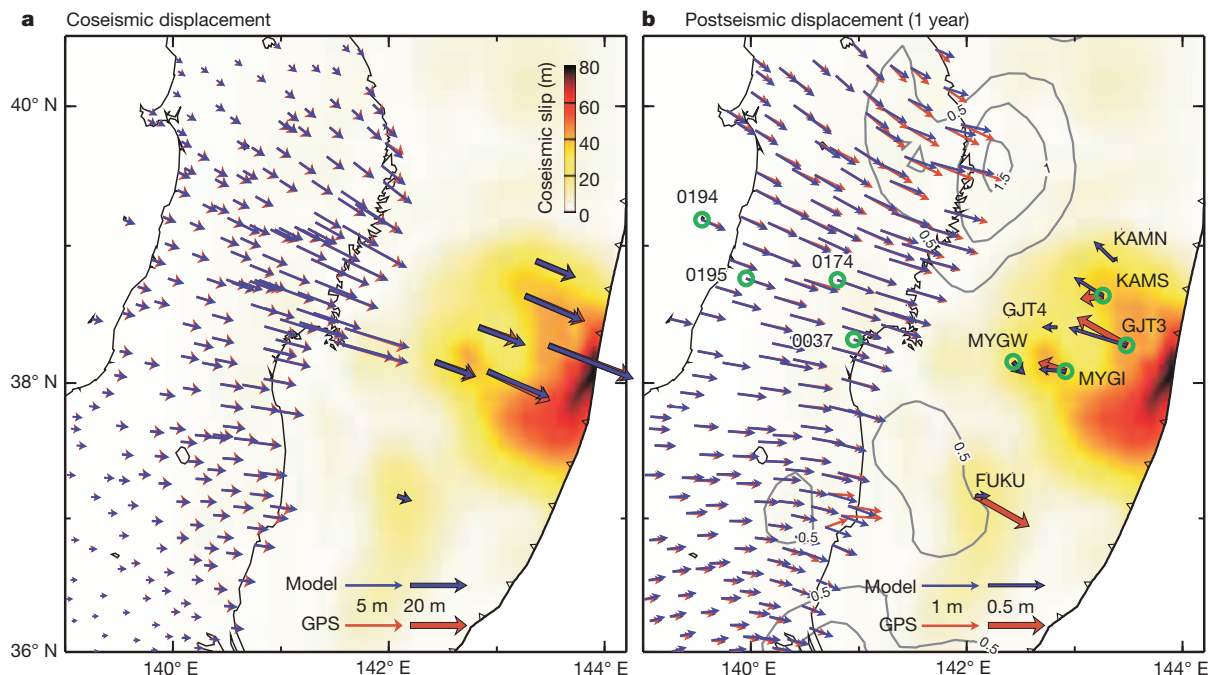
Tianhaozhe Sun<sup>1</sup>, Kelin Wang<sup>1,2</sup>, Takeshi Iinuma<sup>3</sup>, Ryota Hino<sup>3</sup>, Jiangheng He<sup>2</sup>, Hiromi Fujimoto<sup>3</sup>, Motoyuki Kido<sup>3</sup>, Yukihiro Osada<sup>3</sup>, Satoshi Miura<sup>4</sup>, Yusaku Ohta<sup>4</sup> & Yan Hu<sup>5</sup>

After a large subduction earthquake, crustal deformation continues to occur, with a complex pattern of evolution<sup>1</sup>. This postseismic deformation is due primarily to viscoelastic relaxation of stresses induced by the earthquake rupture and continuing slip (afterslip) or relocking of different parts of the fault<sup>2–6</sup>. When postseismic geodetic observations are used to study Earth's rheology and fault behaviour, it is commonly assumed that short-term (a few years) deformation near the rupture zone is caused mainly by afterslip, and that viscoelasticity is important only for longer-term deformation<sup>6,7</sup>. However, it is difficult to test the validity of this assumption against conventional geodetic data. Here we show that new seafloor GPS (Global Positioning System) observations immediately after the great Tohoku-oki earthquake provide unambiguous evidence for the dominant role of viscoelastic relaxation in short-term postseismic deformation. These data reveal fast landward motion of the trench area, opposing the seaward motion of GPS sites on land. Using numerical models of transient viscoelastic mantle rheology, we demonstrate that the landward motion

is a consequence of relaxation of stresses induced by the asymmetric rupture of the thrust earthquake, a process previously unknown because of the lack of near-field observations. Our findings indicate that previous models assuming an elastic Earth will have substantially overestimated afterslip downdip of the rupture zone, and underestimated afterslip updip of the rupture zone; our knowledge of fault friction based on these estimates therefore needs to be revised.

Land-based GPS observations from multiple subduction zones delineate three stages of postseismic deformation following a great megathrust earthquake: wholesale seaward motion, opposing motion of the coastal and inland areas, and wholesale landward motion<sup>1</sup>. This progressive motion reversal contains important information on Earth's viscoelastic rheology and the slip behaviour of subduction megathrusts<sup>1–5</sup>. However, we know surprisingly little about the mechanism of postseismic deformation at the timescale of a few years.

During the Tohoku-oki earthquake, seven seafloor GPS stations operated by the Japan Coast Guard<sup>8</sup> (JCG) and Tohoku University<sup>9</sup> (TU) detected



**Figure 1 | Coseismic and postseismic deformation of the 2011 Tohoku-oki earthquake.** **a**, Coseismic displacements of land (for example, ref. 10) and seafloor<sup>8,9</sup> GPS sites and model predicted displacements based on the fault slip model shown (see Methods). **b**, One-year postseismic displacements of land<sup>15</sup> and seafloor (refs 16 and 17 and Methods) GPS sites and model predicted values based on the viscoelastic model of this work. Seafloor GPS vectors were

obtained through fitting campaign data with logarithmic functions as in Extended Data Fig. 6. Site GJT4 failed shortly after the earthquake. Black contours (m) are the afterslip distribution used in our modelling (see Methods). Observed and model time series at sites marked with a green circle in the main corridor of interest are shown in Fig. 3.

<sup>1</sup>School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia V8P 5C2, Canada. <sup>2</sup>Pacific Geoscience Centre, Geological Survey of Canada, Natural Resources Canada, 9860 West Saanich Road, Sidney, British Columbia V8L 4B2, Canada. <sup>3</sup>International Research Institute of Disaster Science, Tohoku University, Sendai 980-0845, Japan. <sup>4</sup>Research Center for Prediction of Earthquakes and Volcanic Eruptions, Graduate School of Science, Tohoku University, Sendai 980-8578, Japan. <sup>5</sup>Berkeley Seismological Laboratory and Department of Earth and Planetary Sciences, University of California, Berkeley, California, California 94720, USA.

seaward displacements of up to 31 m, much larger than the largest coseismic motion of coastal GPS sites ( $\sim 5$  m; ref. 10) (Fig. 1). Many rupture models, in particular those involving tsunami data and seafloor geodetic observations, feature peak slip exceeding 50 m at rather shallow depths and breaching the trench<sup>11–14</sup>.

After the earthquake, the terrestrial GPS network in northeast Japan continued to show wholesale seaward motion as expected (Fig. 1b). These terrestrial observations can be adequately explained by an afterslip model<sup>15</sup> similar to those developed for most other subduction earthquakes and also based only on terrestrial observations<sup>6</sup>. However, seafloor GPS observations near the trench present a fundamental challenge to the validity of ignoring viscoelastic stress relaxation in short-term postseismic deformation. Whereas some of the seafloor sites also exhibited seaward motion, sites nearest to the peak rupture area immediately reversed their direction from coseismic seaward to postseismic landward (Fig. 1b). These data demonstrate that opposing motion begins immediately after the earthquake, a phenomenon previously unknown because of the lack of seafloor observations. The motion of these sites (Fig. 1b),  $\sim 50$  cm at TU site GJT3 (Extended Data Figs 5 and 6) and  $\sim 20$ –25 cm at JGC sites KAMS and MYGI in the first year (refs 16, 17), is much faster than the subduction rate ( $8.3 \text{ cm yr}^{-1}$ ) at the Japan Trench<sup>18</sup> and thus cannot be explained by the relocking of the subduction fault. Neither can it be explained by afterslip, which would cause the surface to move in the opposite (seaward) direction<sup>6,15</sup>. The effect of poroelastic rebound after the earthquake is far too small to explain the observed motion, even in a model that maximizes such an effect<sup>19</sup>. Therefore, the primary process responsible for this motion must be viscoelastic relaxation<sup>20</sup>.

We explain the immediate landward motion of the trench area, represented by sites GJT3, KAMS and MYGI, as a manifestation of viscoelastic relaxation of stresses induced by the asymmetric rupture of the Tohoku-oki earthquake. We first use a simple two-dimensional (2D) model (Fig. 2a) to elucidate the physical process. This simple model captures the essence of viscoelastic deformation in earthquake cycles<sup>1,3</sup>: the tectonic plates exhibit elastic behaviour, and the asthenospheric mantle deforms elastically at the time of the earthquake but increasingly exhibits viscous behaviour afterwards (Extended Data Fig. 1).

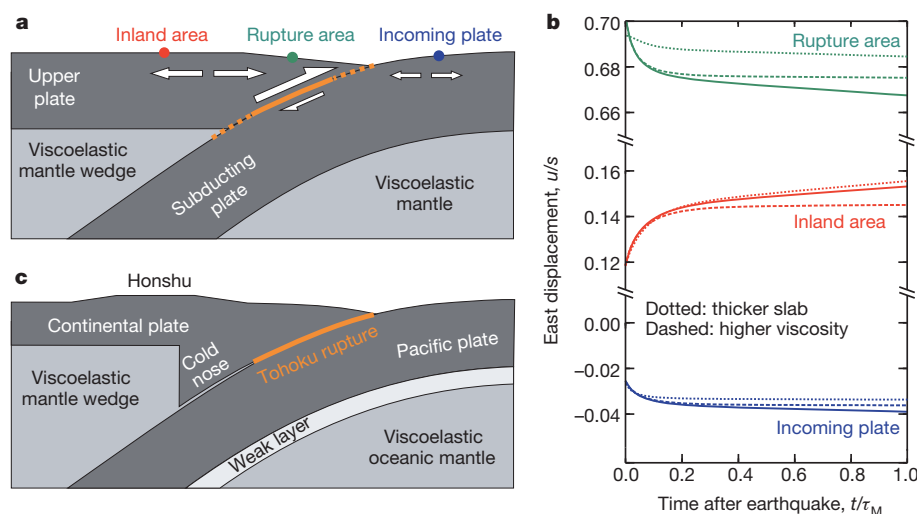
Asymmetric coseismic elastic deformation is a fundamental outcome of any thrust rupture that is not deeply buried. Because of the presence

of the free surface (seafloor), the hanging wall overlying the rupture is less stiff than the foot wall beneath. Consequently, even though the double-couple source mechanism is symmetric, the hanging wall undergoes greater coseismic motion than does the foot wall. In subduction earthquakes, the asymmetry is very pronounced owing to the shallow dip and depth of the megathrust (Fig. 2a). The maximum seaward motion of the upper plate is larger than the maximum landward motion of the incoming plate often by an order of magnitude. Systematic spatial variations in rock rigidity or plastic yielding in parts of the system can only slightly modify the relative magnitude of the motion but cannot reduce the asymmetry in any substantive fashion.

The asymmetric rupture induces greater tension in the upper plate than in the incoming plate (Fig. 2a). The stress asymmetry around the rupture zone is accompanied by heterogeneous incremental stresses in the rest of the coseismically elastic system that, for static deformation, balance the net force and torque. As the underlying mantle undergoes viscoelastic relaxation after the earthquake<sup>1,3</sup>, the greater tension in the upper plate pulls the trench area landward (Fig. 2b). The site in the rupture area reverses its direction of motion immediately after the earthquake in all the models, irrespective of the vastly different parameters used. For example, a very thick subducting plate or highly viscous mantle can slow down the motion but cannot prevent it from occurring (Fig. 2b).

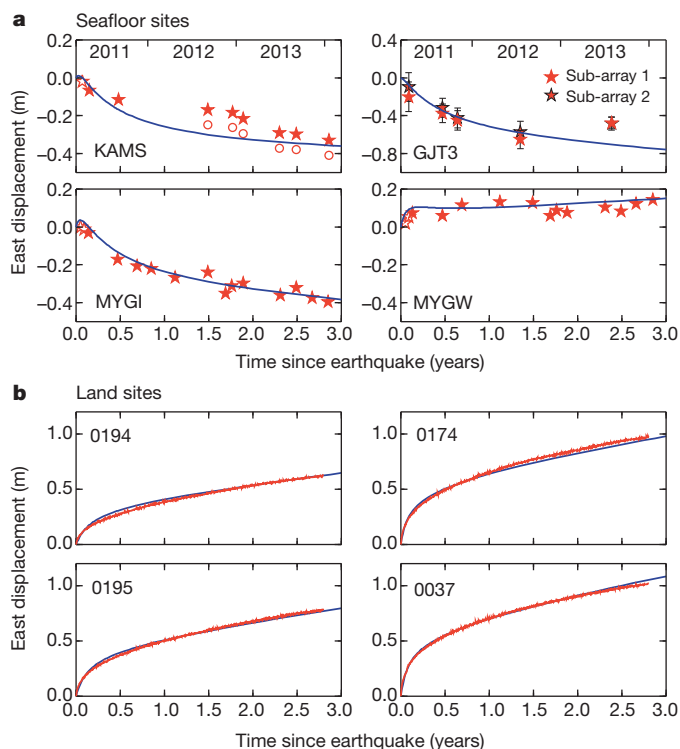
In the real Earth, the only process that may offset or even reverse this motion in limited areas is fast afterslip, especially at very shallow depths such as that observed after the 2005 moment magnitude  $M_w = 8.7$  Sumatra earthquake<sup>7</sup>. At the Japan Trench, the very fast seaward motion of JCG site FUKU, outside the main rupture area (Fig. 1b), is undoubtedly caused by shallow afterslip. We think that the lack of landward motion of JCG site KAMN is probably because the motion was offset by local afterslip, an issue that we do not have adequate information to explore in our modelling.

To apply the conceptual model illustrated in Fig. 2a and b to the seafloor GPS observations after the Tohoku-oki earthquake, we developed a three-dimensional (3D) spherical-Earth finite element model (see Methods) involving the Burgers mantle rheology (Extended Data Fig. 1) and the actual fault geometry (Extended Data Fig. 2). Our main region of interest is the broad margin-normal corridor including the peak rupture area and sites GJT3, MYGI, KAMS and MYGW. To focus on the first-order



**Figure 2 | Numerical models of short-term viscoelastic relaxation.** **a**, A 2D generic subduction earthquake model used to illustrate the consequence of asymmetric rupture. Fault slip,  $s$ , is denoted by a solid orange line, and tapers to zero over dashed portions. Greater tensile stress is coseismically induced in the upper plate than in the incoming plate (diverging arrows). **b**, Horizontal coseismic ( $t = 0$ ) and postseismic ( $t > 0$ ) displacements ( $u$ ) of the three colour coded sites in Fig. 2a in response to the earthquake.  $\tau_M$  is the Maxwell time of the mantle wedge (Extended Data Fig. 1). Solid lines show results based on a

model with 30-km-thick upper and lower plates and a rigidity and viscosity structure similar to previous subduction earthquake cycle models<sup>1,3,23</sup> and identical to model B in Extended Data Table 1. In the ‘thicker slab’ model, the lower plate is twice as thick; in the ‘higher viscosity’ model, the steady-state mantle wedge viscosity is an order of magnitude higher ( $10^{20}$  Pa s). **c**, Schematic illustration of the structure of the 3D model for the 2011 Tohoku-oki earthquake, with results shown in Figs 1 and 3.



**Figure 3 | Observed (red) and model-predicted (blue) time series of the east component of postseismic displacements.** The locations of the GPS sites are shown in Fig. 1b. **a**, Seafloor sites. For TU site GJT3, error bars (standard error) are based on error analysis, and sub-arrays are formed by different combinations of seafloor transponders, both as explained in Methods. For the JCG sites<sup>16,17</sup>, error estimates were not provided but are estimated to be smaller than those of GJT3 (see Methods) except for the first one or two less reliable measurements at each site (open stars). Circles for KAMS represent position data after a manual correction for an assumed delayed local afterslip during 2012. The one-year vector for this site shown in Fig. 1b is based on the corrected data. **b**, Randomly selected land sites in the main corridor of interest. Other sites in this corridor have similar results.

physical process, we purposely simplified the model by using uniform material properties for each of the major structural units (model A in Extended Data Table 1). Elastic modulus values are the same as in ref. 1 except for those required by the transient rheology (Extended Data Fig. 1), for which larger values better reproduce postseismic motion of all the land GPS sites in the first few weeks. The most seaward part of the mantle wedge overlying the shallower-than-70-km part of the slab is an elastic ‘cold nose’ (Fig. 2c), representing the stagnant and cold part of the mantle wedge<sup>21</sup> and consistent with the results of seismic tomography in this region<sup>22</sup>. Between the cold nose and the slab, the plate interface changes from a distinct fault at shallow depths to a thin viscoelastic shear zone at greater depths (see Methods). Differently from previous models, we included a weak layer (Extended Data Table 1) below the oceanic plate, approximately accounting for the recently but widely reported mechanical decoupling of the oceanic lithosphere from the underlying mantle material (see Methods). All the viscosity values were optimized to fit observations via a trial-and-error approach.

We used a coseismic rupture model slightly modified (see Methods) from ref. 12 (Fig. 1a). Our tests show that different choices of coseismic slip models<sup>14</sup> may lead to slightly different estimates of viscosity values but do not change the physical process demonstrated by the model. Because postseismic GPS observations reflect both viscoelastic relaxation and afterslip, we must consider both processes to allow meaningful comparisons with data<sup>23</sup>. We opted to revise the afterslip model of ref. 15 and combine it into our 3D viscoelastic model in a trial-and-error fashion. The introduction of viscoelasticity as required by seafloor observations greatly reduced the amount of afterslip required to explain the

land GPS data. The afterslip values shown in Fig. 1b have been reduced from those of ref. 15 by as much as 95% directly down-dip of the main rupture area and by about 30% farther away (see Methods). Our model does not include shallow and/or trench-breaching afterslip and therefore is not designed to explain the motion of site FUKU (see Methods). If significant shallow afterslip did occur in our main corridor of interest, the landward motion of seafloor GPS sites due to viscoelastic relaxation should be even faster than shown in Fig. 1b, further strengthening the main argument of this Letter.

Our 3D model adequately explains the spatial (Fig. 1b) and temporal (Fig. 3) patterns of postseismic deformation. Even in areas away from the main corridor of interest, the model fits GPS observations to a considerable degree of fidelity. Second-order temporal variations in the GPS time series, such as the brief slowing down of KAMS during 2012 and the motion reversal of GJT3 in 2013, may be due to local adjustment of the megathrust (delayed afterslip) and cannot be explained by viscoelastic relaxation (Fig. 3a). Steady-state viscosities in this model (model A in Extended Data Table 1) are lower than in previous models<sup>1</sup> that were based mostly on longer-term postseismic and interseismic observations (see also Extended Data Figs 3 and 4). The reason is most probably that transient mantle rheology is more complex than described by the Burgers model (Extended Data Fig. 1) and our steady-state viscosity based on the ~3 years of postseismic observations may still be affected by transient creep.

Our numerous testing runs using both 2D and 3D models (not all displayed here) show that landward trench motion does not occur in any purely elastic model but always occurs in viscoelastic models irrespective of the details of the viscoelastic mantle rheology, afterslip and model structure. Therefore, in elastic models for any large subduction earthquakes, afterslip down-dip of the rupture zone will have been overestimated, and afterslip at shallower depths, if present and resolvable by observations<sup>7</sup>, will have been underestimated. Reassessing afterslip using viscoelastic models will lead to a revision of our knowledge of the slip behaviour and physics of subduction megathrusts.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 July; accepted 20 August 2014.

Published online 17 September 2014.

- Wang, K., Hu, Y. & He, J. Deformation cycles of subduction earthquakes in a viscoelastic Earth. *Nature* **484**, 327–332 (2012).
- Hu, Y., Wang, K., He, J., Klotz, J. & Khazaradze, G. Three-dimensional viscoelastic finite element model for post-seismic deformation of the great 1960 Chile earthquake. *J. Geophys. Res.* **109**, B12403 (2004).
- Pollitz, F. F., Bürgmann, R. & Banerjee, P. Post-seismic relaxation following the great 2004 Sumatra-Andaman earthquake on a compressible self-gravitating Earth. *Geophys. J. Int.* **167**, 397–420 (2006).
- Suito, H. & Freymueller, J. T. A viscoelastic and afterslip postseismic deformation model for the 1964 Alaska earthquake. *J. Geophys. Res.* **114**, B11404 (2009).
- Kogan, M. G. et al. Rapid postseismic relaxation after the great 2006–2007 Kuril earthquakes from GPS observations in 2007–2011. *J. Geophys. Res.* **118**, 3691–3706 (2013).
- Pritchard, M. E. & Simons, M. An aseismic slip pulse in northern Chile and along-strike variations in seismogenic behavior. *J. Geophys. Res.* **111**, B08405 (2006).
- Hsu, Y.-J. et al. Frictional afterslip following the 2005 Nias-Simeulue earthquake, Sumatra. *Science* **312**, 1921–1926 (2006).
- Sato, M. et al. Displacement above the hypocenter of the 2011 Tohoku-Oki earthquake. *Science* **332**, 1395 (2011).
- Kido, M., Osada, Y., Fujimoto, H., Hino, R. & Ito, Y. Trench-normal variation in observed seafloor displacements associated with the 2011 Tohoku-Oki earthquake. *Geophys. Res. Lett.* **38**, L24303 (2011).
- Ozawa, S. et al. Coseismic and postseismic slip of the 2011 magnitude-9 Tohoku-Oki earthquake. *Nature* **475**, 373–376 (2011).
- Fujii, Y., Satake, K., Sakai, S., Shinohara, M. & Kanazawa, T. Tsunami source of the 2011 off the Pacific coast of Tohoku earthquake. *Earth Planets Space* **63**, 815–820 (2011).
- Iinuma, T. et al. Coseismic slip distribution of the 2011 off the Pacific Coast of Tohoku Earthquake (M9.0) refined by means of seafloor geodetic data. *J. Geophys. Res.* **117**, B07409 (2012).
- Shao, G., Chen, J. & Archuleta, R. Quality of earthquake source models constrained by teleseismic waves: using the 2011 M9 Tohoku-oki earthquake as an example. (Poster 93, presented at Incorporated Research Institutions for Seismology



- Workshop, Boise, Idaho, 13–15 June, 2012); available at [http://www.iris.edu/hq/iris\\_workshop2012/sci/0115.html](http://www.iris.edu/hq/iris_workshop2012/sci/0115.html).
14. Tajima, F., Mori, J. & Kennett, B. L. N. A review of the 2011 Tohoku-oki earthquake (Mw 9.0): large-scale rupture across heterogeneous plate coupling. *Tectonophysics* **586**, 15–34 (2013).
  15. Ozawa, S. *et al.* Preceding, coseismic, and postseismic slips of the 2011 Tohoku earthquake, Japan. *J. Geophys. Res.* **117**, B07404 (2012).
  16. Japan Coast Guard & Tohoku University. Seafloor movements observed by seafloor geodetic observations after the 2011 off the Pacific coast of Tohoku earthquake. *Rep. Coord. Committee Earthquake Prediction* **90**, 3–4 (2013).
  17. Watanabe, S. *et al.* Evidence of viscoelastic deformation following the 2011 Tohoku-oki earthquake revealed from seafloor geodetic observation. *Geophys. Res. Lett.* (in the press); preprint at <http://onlinelibrary.wiley.com/doi/10.1002/2014GL061134/abstract>.
  18. DeMets, C., Gordon, R. G. & Argus, D. F. Geologically current plate motions. *Geophys. J. Int.* **181**, 1–80 (2010).
  19. Hu, Y., Burgmann, R., Freymueller, J. F., Banerjee, P. & Wang, K. Contributions of poroelastic rebound and a weak volcanic arc to the postseismic deformation of the 2011 Tohoku earthquake. *Earth Planets Space* **66**, 106 (2014).
  20. Sun, T. *et al.* Viscoelastic landward motion of the trench area following a subduction earthquake. Abstr. G14A–08 (Fall Meeting, AGU, San Francisco, 9–13 December, 2013); available at <http://adsabs.harvard.edu/abs/2013AGUFM.G14A..08S>.
  21. Wada, I. & Wang, K. Common depth of decoupling between the subducting slab and mantle wedge: reconciling diversity and uniformity of subduction zones. *Geochem. Geophys. Geosyst.* **10**, Q10009 (2009).
  22. Yamamoto, Y., Hino, R. & Shinohara, M. Mantle wedge structure in the Miyagi Prefecture forearc region, central northeastern Japan arc, and its relation to corner-flow pattern and interplate coupling. *J. Geophys. Res.* **116**, B10310 (2011).
  23. Hu, Y. & Wang, K. Spherical-Earth finite element model of short-term postseismic deformation following the 2004 Sumatra earthquake. *J. Geophys. Res.* **117**, B05404 (2012).

**Acknowledgements** We thank the Japan Coast Guard for making available digital values of published data. Comments from M. Sato improved the manuscript. K.W. was supported by Geological Survey of Canada core funding and a Natural Sciences and Engineering Research Council of Canada Discovery Grant through the University of Victoria. T.S. was supported by a University of Victoria PhD Fellowship and a Howard E. Petch Scholarship. The Tohoku University seafloor observation study was supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan under its Earthquake and Volcano Hazards Observation and Research Program. This is Geological Survey of Canada contribution 20140167.

**Author Contributions** T.S. carried out the numerical modelling. K.W. designed the study. K.W. and T.S. together did most of the writing. T.J. processed land GPS data. R.H., H.F., M.K., Y. Osada, S.M. and Y. Ohta collected and processed GJT3 seafloor GPS data. J.H. wrote the modelling code and contributed to the modelling. Y.H. constructed fault geometry and initiated the modelling.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.W. ([kwang@nrcan.gc.ca](mailto:kwang@nrcan.gc.ca)).

## METHODS

**Finite element model.** We assume that the mantle obeys the bi-viscous Burgers rheology<sup>1,3</sup>. The Kelvin solid of viscosity  $\eta_K$  and rigidity  $\mu_K$  and the Maxwell fluid of viscosity  $\eta_M$  and rigidity  $\mu_M$  in the Burgers body (Extended Data Fig. 1) are the simplest parameterizations of the transient and steady-state rheology, respectively<sup>24</sup>. The characteristic timescales of the transient and steady-state rheology are thus represented by the Kelvin relaxation time  $\tau_K = \eta_K/\mu_K$  and Maxwell relaxation time  $\tau_M = \eta_M/\mu_M$ , respectively. Note that  $\mu_K$  is not a real physical property but a parameter introduced to control the initial rate of transient creep of mantle material without invoking more parameters.

Secular mantle wedge flow maintains high temperatures in the arc and back arc region<sup>21</sup>. The different thermal states of the two sides can result in not only different thicknesses of the elastic plates, but also differences in the viscosities of the mantle below. Following the arguments of ref. 1, we required the viscosities of the mantle wedge to be about one order of magnitude lower than those of the oceanic mantle (Extended Data Table 1).

We used the spherical-Earth finite-element code PGCviscl-3D developed by J.H. The code uses 27-node isoparametric elements throughout the model domain. The effect of gravitation is incorporated using the stress-advection approach<sup>25</sup>. Coseismic rupture and afterslip are simulated using the split-node method<sup>26</sup>. Time ( $t$ ) integration is performed using a fully implicit algorithm, with time steps no greater than  $0.01\tau_K$  for  $t < \tau_K$  and no greater than  $0.01\tau_M$  for  $t < 0.5\tau_M$ . The parallel code has been extensively benchmarked against analytical deformation solutions for elastic, Maxwell and Burgers materials and applied to subduction zone earthquake cycle modelling<sup>1,23</sup>.

The central part of the element mesh for the Tohoku-oki model is shown in Extended Data Fig. 2. The subduction fault geometry was constrained by earthquake relocation results and seismic reflection profiles<sup>27–29</sup> and is similar to what was used in ref. 12. We accounted for the presence of a cold and stagnant nose of the mantle wedge<sup>21,22</sup> and its sharp landward termination<sup>30</sup> by adding a triangular region to the elastic upper plate in the forearc (Fig. 2c and Extended Data Fig. 2).

Studies of fault processes indicate that the distinction between shear along a thin fault plane and within a broader shear zone becomes blurry at large depths<sup>31</sup>. Much of the afterslip is actually shear deformation that gradually spreads over a shear zone that thickens with increasing depth. In our model, between the elastic cold nose and the elastic slab is a layer of viscoelastic mantle material that thickens with increasing depth (Fig. 2c and Extended Data Fig. 2). This layer approximates the deeper fault zone to a depth of 70 km. Deeper than 70–80 km, the mantle wedge is fully coupled with the slab, that is, there is no longer a fault zone that accommodates localized shear such as afterslip<sup>21</sup>.

Recent studies suggest mechanical decoupling at the lithosphere–asthenosphere boundary (LAB)<sup>32–34</sup>, due to the presence of either fluids<sup>35</sup> or partial melts<sup>36,37</sup>. We thus introduced a thin layer of low viscosity underlying the elastic oceanic plate to approximate this effect (Fig. 2c and Extended Data Table 1). This approximate LAB layer decreases the ratio of vertical to horizontal postseismic displacements at the seafloor. Compared to models without this layer, our model predicts smaller postseismic subsidence in the rupture area and is generally more consistent with observations. However, because of the much larger errors in observed vertical deformation, we did not try to fit the vertical data precisely.

**Assigning coseismic slip and afterslip.** In ref. 12, terrestrial and seafloor GPS and ocean bottom pressure data were inverted using a model of a planar fault to determine the coseismic slip distribution. We mapped the slip vectors onto our 3D curved fault surface. The original slip model used a straight line to represent the trench, resulting in a gap between the model rupture zone and the actual curved trench or some slip seaward of the trench. We filled the artificial trench gap by extrapolating slip values from the model rupture zone (Fig. 1a), and the additional slip resulted in a larger seismic moment and surface displacements. We scaled the fault slip to 92% of its original values in order to match the GPS observations (Fig. 1a). We have developed postseismic deformation models using other published rupture models. Different coseismic slip distributions require slightly different mantle viscosity values in order to fit the GPS data, but all lead to the same main conclusions.

The afterslip model shown in Fig. 1b (contours) was revised from the model of afterslip 8 months after the earthquake developed in ref. 15. Because the model of ref. 15 assumed a purely elastic Earth, postseismic deformation caused by viscoelastic relaxation was also attributed to afterslip, resulting in over-estimated afterslip. Therefore, we scaled down the afterslip values when assigning them to our finite element mesh. Unlike the uniform scaling ratio used for coseismic slip, we needed to use a smoothly variable function for the afterslip scaling. With trial-and-error, the scaling factor was determined to be  $\sim 0.05$  down-dip of the main rupture zone at  $\sim 60$ –70 km depth,  $\sim 0.35$  to the north of the main rupture zone, and  $\sim 0.7$  to the south of the main rupture zone. For the temporal evolution of the afterslip, we used the power-law function reported in ref. 23 with a characteristic timescale of 1.5 years.

Our model does not include any shallow afterslip near or breaching the trench. For our main corridor of interest, the assumption of no shallow afterslip is supported by the fact that a postseismic thermal-sensor monitoring string deployed in a near-trench borehole was retrieved intact<sup>38</sup>, indicating no trench-breaching afterslip at this site during the monitoring period (16–24 months after the Tohoku-oki earthquake). If there is significant shallow afterslip before the monitoring period or in other parts of our main corridor of interest, the actual landward motion of sites GJT3, MYGI and KAMS due to viscoelastic relaxation should be even faster than shown by the GPS data. For this reason, our model represents a minimum estimate of the effect of viscoelastic relaxation.

**Model using the viscosity values of ref. 1.** Testing model B (Extended Data Table 1) shows why we cannot use the viscosity structure and values used in ref. 1. A mantle wedge Maxwell viscosity of  $10^{19}$  Pa s was used in ref. 1 for a study of longer-term postseismic deformation. If the same value is used in our model, it is possible to explain cumulative GPS displacements observed at a specific time (Extended Data Fig. 3) but very difficult to explain the time-dependent evolution of the deformation field (Extended Data Fig. 4).

**Seafloor/acoustic observation at GJT3.** GJT3 operated by Tohoku University is the most important seafloor GPS site in this study because it is the nearest to the trench. The basic concept of the GPS/acoustic technique used by Tohoku University to make seafloor geodetic measurements was developed originally by the Scripps Institution of Oceanography<sup>39,40</sup>. The technique measures the horizontal displacement of the virtual seafloor benchmark, the centre of an array of at least three seafloor precision transducers (PXP), by repeated surveys using a sea surface platform equipped with GPS antennas and an acoustic transducer<sup>41</sup>. Two survey methods can be used. In the fixed-point survey method, routinely used by Tohoku University, the surface platform is placed above the centre of the PXP array. If the array geometry does not change with time, the fixed-point survey method can be used to monitor the horizontal motion of the virtual benchmark. In the moving survey method, routinely used by JCG<sup>42</sup>, the platform moves around each individual PXP to determine its absolute position. This procedure is more robust because no assumptions on PXP array geometry are required, but it is very time consuming.

Given precise position of the surface platform, two-way travel times between the platform and the PXPs, and knowledge of temporal variations in underwater sound speed, the horizontal position of the array is determined by simultaneous ranging of a single acoustic ping to all the transponders<sup>43</sup>. If the sound speed structure is horizontally stratified, this method is expected to give reliable estimates of the array position. However, temporal changes and three-dimensional heterogeneities of the sound speed structure often cause the position measurements to fluctuate. During a campaign, we take an ensemble mean of many measurements to estimate the array position, such that much of the effects of the sound speed anomalies are averaged out.

Within the first two years after the Tohoku-oki earthquake, Tohoku University conducted four campaign surveys at GPS/acoustic station GJT3, located above the main rupture area (Fig. 1). The first measurement, made in April 2011, showed a displacement of about 31 m due mainly to coseismic motion<sup>9</sup>. This and the two subsequent surveys in 2012 used only the fixed-point method because of limited ship time allocation. In 2013, we used the moving survey method to reassess the array geometry at this site while continuing to use the fixed-point method to determine the position of the virtual seafloor benchmark. The moving survey results indicated that the PXP array geometry had changed, most likely during the earthquake. Given the proximity of the site to the peak rupture area (Fig. 1), this finding is not surprising. For the very large coseismic displacement<sup>9</sup>, errors due to incorrectly assuming rigid array geometry are negligibly small. For the much smaller postseismic displacements, however, this assumption leads to significant errors. We conclude that the postseismic motion of GJT3 based on the fixed-point survey results of 2012 alone<sup>44</sup> had yielded an incorrect direction of motion. There is no obvious reason why the array geometry would have suffered further significant distortion after the earthquake. Therefore, in the present study, we reprocessed all the postseismic data using the PXP geometry newly determined in 2013.

The JCG array positions shown in Figs 1 and 3 were determined by averaging the positions of individual PXPs. Without invoking fixed-point survey, a large amount of ship time is required in order to minimize errors caused by the uncertainties of individual PXP locations. However, because no assumptions about array geometry are involved, the locations of the seafloor benchmark estimated by JCG are minimally affected by potential coseismic distortion of array geometry.

Regardless of the survey method, uncertainties in the position of individual PXPs can be a source of error in estimating the array position. When a fixed-point survey is made near the array centre, uncertainties in PXP positions do not affect the estimation of the array position. However, the estimation error rapidly increases with the offset of the surface platform from the array centre. Keeping the platform at the centre was especially difficult during the survey in April 2011 when large amounts

of tsunami debris drifted around the site and prevented the research vessel from staying at the optimum location.

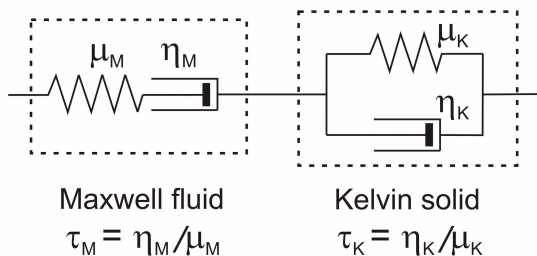
Another factor we have to take into account is site displacement caused by nearby major aftershocks. An  $M_w$  7 intraslab earthquake occurred on 10 July 2011, and had a strike-slip mechanism. The epicentre is only about 20 km from GJT3 and induced coseismic displacement that cannot be ignored. Its fault location was defined by the aftershock distribution precisely determined with an ocean bottom seismic network<sup>45</sup>, and its slip model was estimated from near-field tsunami waveforms<sup>46</sup>. Using this information, we estimated that the displacement of GJT3 due to this event was 2.0 cm westward and 6.6 cm southward.

Extended Data Fig. 5 shows the PXP array configuration at GJT3. Excluding PXPs EJ16 and EJ23, which were installed for testing purposes, the PXPs form an equilateral triangle with side length of  $\sim 2.5$  km. Two PXPs (EJ15 and EJ22) are collocated at one of the apexes. Since the array position can be determined using the fixed-point observation with three seafloor PXPs, we can have two different sub-arrays: sub-array 1, composed of PXPs EJ15, EJ12 and EJ13, and sub-array 2, composed of PXPs EJ22, EJ12 and EJ13. Extended Data Fig. 6 shows the time series of the array positions of the two sub-arrays after a correction for the effect of the 2011  $M_w$  7 earthquake discussed above. Position error in each campaign is estimated from the root-mean-squares of position measurements around the mean position and uncertainties in PXP positioning. Here, we assumed that the PXP positions determined by the moving survey method contain 1 m uncertainties, based on uncertainties in the sound speed of the order of 0.01% and the slant ranges from the surface platform to PXPs at  $\sim 4,000$  m. Consistency between the two sub-arrays suggests the robustness of the results.

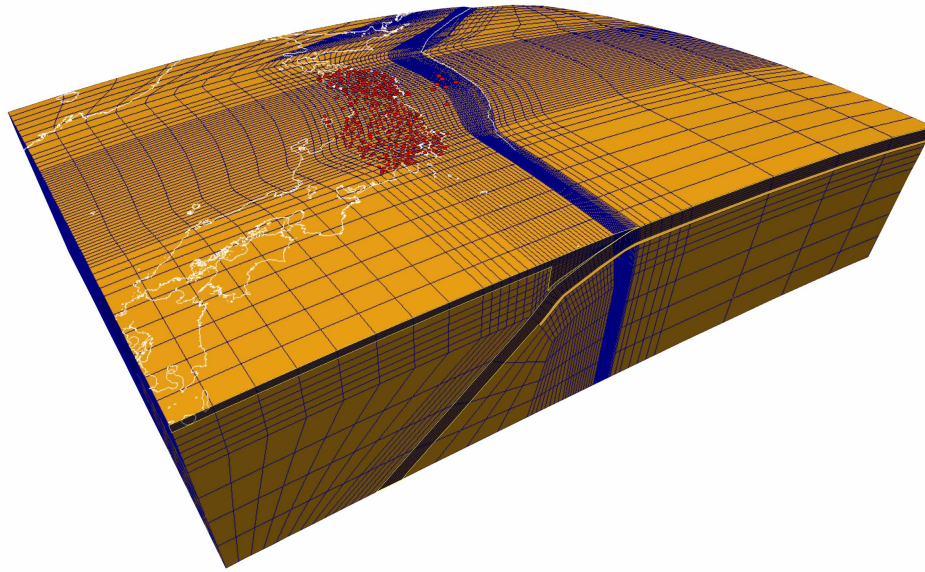
24. Peltier, W. R., Wu, P. & Yuen, D. A. in *Anelasticity in the Earth* (eds Stacey, F. D., Paterson, M. S. & Nicolas, A.) 59–77 (Geodynamics Ser. Vol. 4, American Geophysical Union, 1981).
25. Peltier, W. R. The impulse response of a Maxwell Earth. *Rev. Geophys. Space Phys.* **12**, 649–668 (1974).
26. Melosh, H. J. & Raefsky, A. A simple and efficient method for introducing faults into finite element computations. *Bull. Seismol. Soc. Am.* **71**, 1391–1400 (1981).
27. Nakajima, J. & Hasegawa, A. Anomalous low-velocity zone and linear alignment of seismicity along it in the subducted Pacific slab beneath Kanto, Japan: reactivation of subducted fracture zone? *Geophys. Res. Lett.* **33**, L16309 (2006).
28. Kita, S., Okada, T., Hasegawa, A., Nakajima, J. & Matsuzawa, T. Anomalous deepening of a seismic belt in the upper-plane of the double seismic zone in the Pacific slab. *Earth Planet. Sci. Lett.* **290**, 415–426 (2010).
29. Zhao, D., Wang, Z., Umino, N. & Hasegawa, A. Mapping the mantle wedge and interplate thrust zone of the northeast Japan arc. *Tectonophysics* **467**, 89–106 (2009).
30. Wada, I., Rychert, C. A. & Wang, K. Sharp thermal transition in the forearc mantle wedge as a consequence of nonlinear mantle wedge flow. *Geophys. Res. Lett.* **38**, L13308 (2011).
31. Noda, H. & Shimamoto, T. Transient behavior and stability analyses of halite shear zones with an empirical rate-and-state friction to flow law. *J. Struct. Geol.* **38**, 234–242 (2012).
32. Kawakatsu, H. *et al.* Seismic evidence for sharp lithosphere-asthenosphere boundaries of oceanic plates. *Science* **324**, 499–502 (2009).
33. Rychert, C. A. & Shearer, P. M. A global view of the lithosphere-asthenosphere boundary. *Science* **324**, 495–498 (2009).
34. Fischer, K. M., Ford, H. A., Abt, D. L. & Rychert, C. A. The lithosphere-asthenosphere boundary. *Annu. Rev. Earth Planet. Sci.* **38**, 551–575 (2010).
35. Karato, S. On the origin of the asthenosphere. *Earth Planet. Sci. Lett.* **321–322**, 95–103 (2012).
36. Sakamaki, T. *et al.* Ponded melt at the boundary between the lithosphere and asthenosphere. *Nature Geosci.* **6**, 1041–1044 (2013).
37. Schmerr, N. The gutenberge discontinuity: melt at the lithosphere-asthenosphere boundary. *Science* **335**, 1480–1483 (2012).
38. Fulton, P. M. *et al.* Low coseismic friction on the Tohoku-oki fault determined from temperature measurements. *Science* **342**, 1214–1217 (2013).
39. Spiess, F. N. Suboceanic geodetic measurements. *IEEE Trans. Geosci. Remote Sens.* **GE-23**, 502–510 (1985).
40. Fujimoto, H. Seafloor geodetic approaches to subduction thrust earthquakes. *Monogr. Environ. Earth Planets* **2**, 23–63 (2014).
41. Kido, M. *et al.* Seafloor displacement at Kumano-nada caused by the 2004 off Kii Peninsula earthquakes, detected through repeated GPS/acoustic surveys. *Earth Planets Space* **58**, 911–915 (2006).
42. Sato, M. *et al.* Interplate coupling off northeastern Japan before the 2011 Tohoku-oki earthquake, inferred from seafloor geodetic data. *J. Geophys. Res.* **118**, 3860–3869 (2013).
43. Kido, M., Osada, Y. & Fujimoto, H. Temporal variation of sound speed in ocean: a comparison between GPS/acoustic and in situ measurements. *Earth Planets Space* **60**, 229–234 (2008).
44. Osada, Y. *et al.* Seafloor crustal movement observed off Miyagi after the 2011 Tohoku earthquake using GPS-acoustic observation system. Abstr. T13F-2693 (Fall Meeting, AGU, 2012); available at <http://adsabs.harvard.edu/abs/2012AGUFM.T13F2693O>.
45. Obana, K. *et al.* Aftershocks near the updip end of the 2011 Tohoku-Oki earthquake. *Earth Planet. Sci. Lett.* **382**, 111–116 (2013).
46. Kubota, T. *et al.* Source models of M-7 class earthquakes in the rupture area of the 2011 Tohoku-Oki earthquake by near-field tsunami modeling. Abstr. T13B-2594 (Fall Meeting, AGU, 2012); available at <http://adsabs.harvard.edu/abs/2012AGUFM.T13B2594K>.



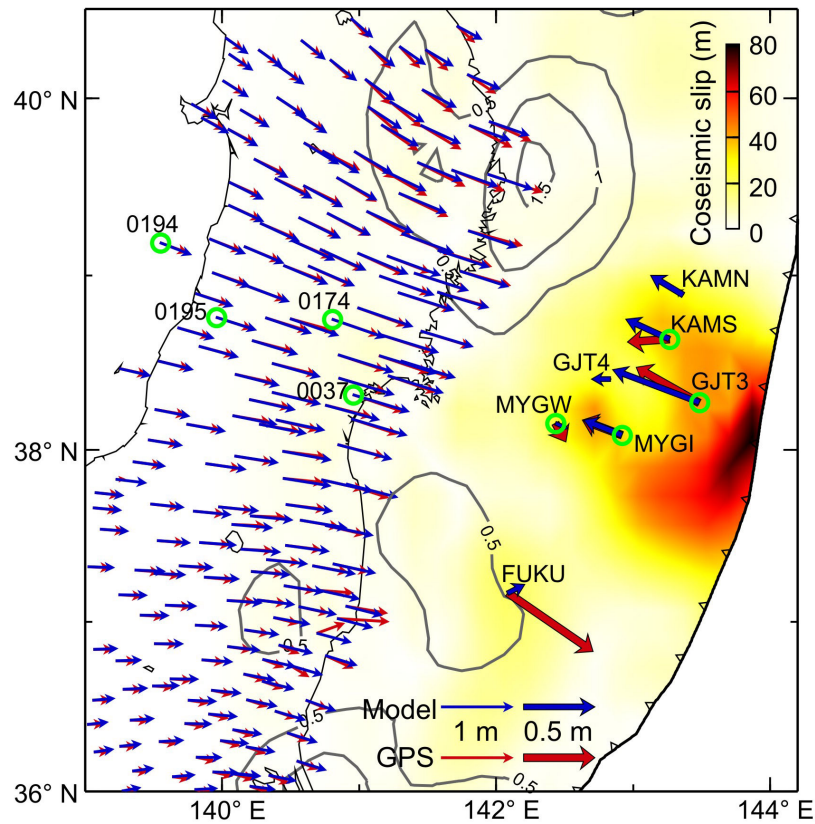
## Burgers rheology



**Extended Data Figure 1 | Illustration of the Burgers rheology used in this work.** The Burgers rheology is represented by a serial connection of a Maxwell fluid of viscosity  $\eta_M$  and rigidity  $\mu_M$  and a Kelvin solid of viscosity  $\eta_K$  and rigidity  $\mu_K$ .  $\tau_M$  and  $\tau_K$  are Maxwell and Kelvin relaxation times, respectively.

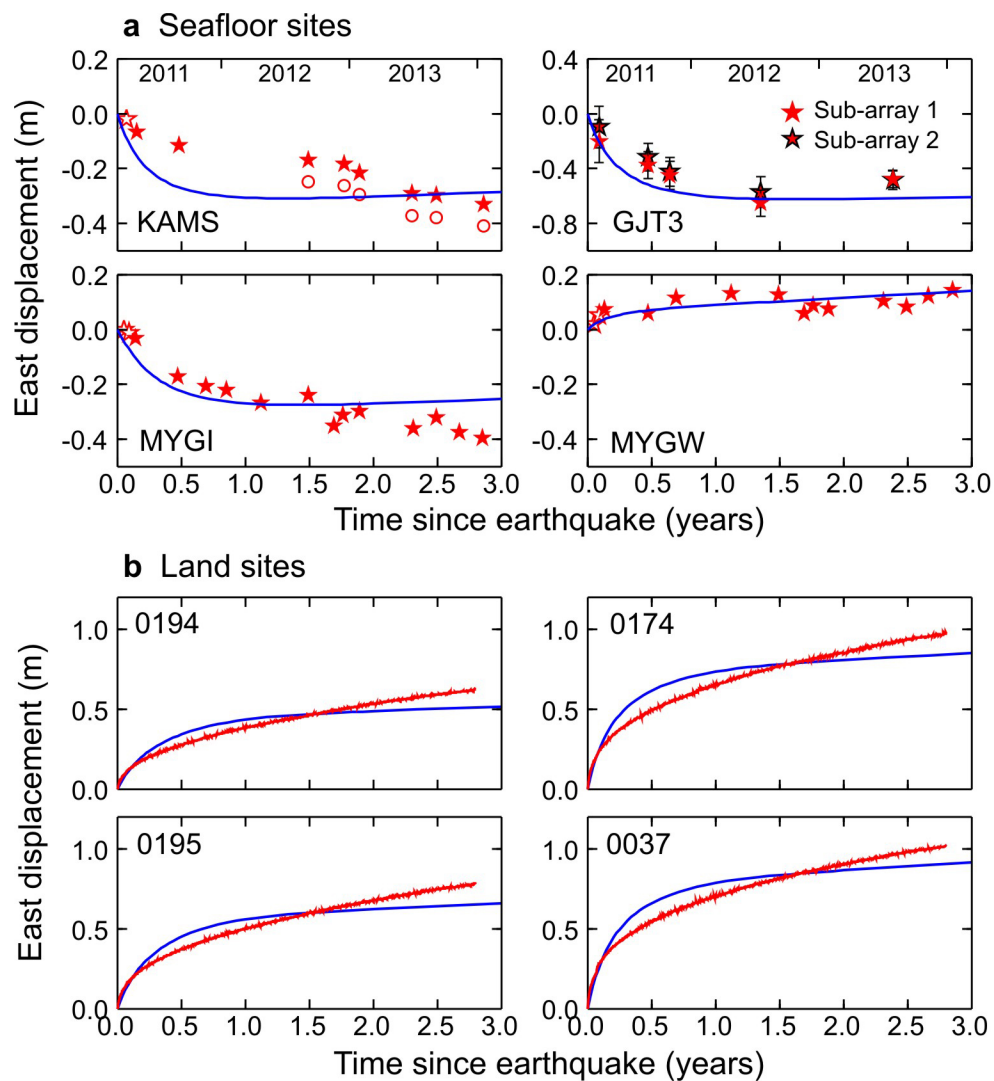


**Extended Data Figure 2 | Central part of the finite element mesh for modelling deformation associated with the Tohoku-oki earthquake.** Darker layers represent elastic plates. The LAB layer is highlighted in yellow. Structural details are shown in Fig. 2c. GPS sites used to constrain the model in this work are shown in red. Elements near the trench are too fine to be discerned at this plotting scale and hence collectively appear as a blue region.

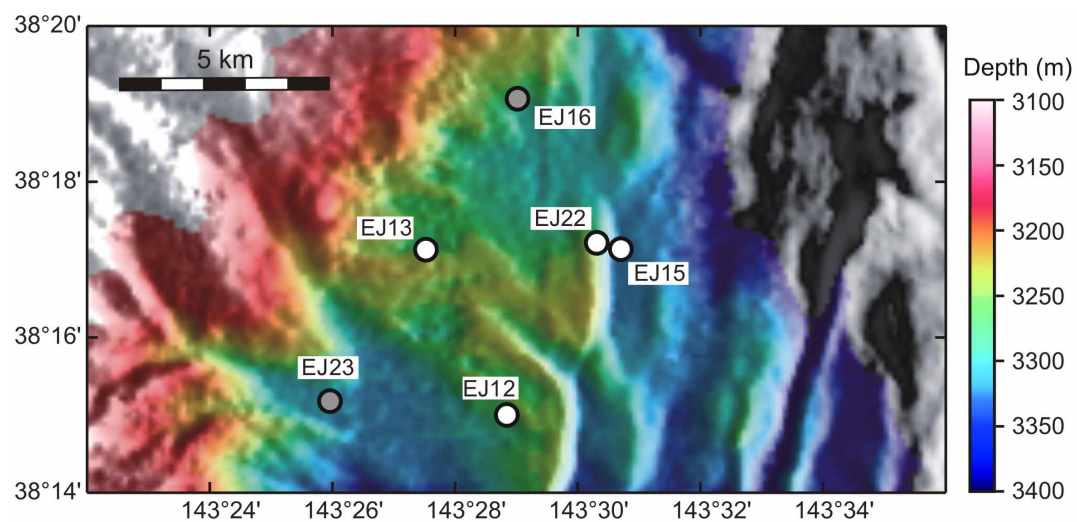


**Extended Data Figure 3 |** Postseismic (1 year) deformation results of model B in Extended Data Table 1. Otherwise the figure is the same as Fig. 1b. Time series at sites marked with a green circle are shown in Extended Data Fig. 4.

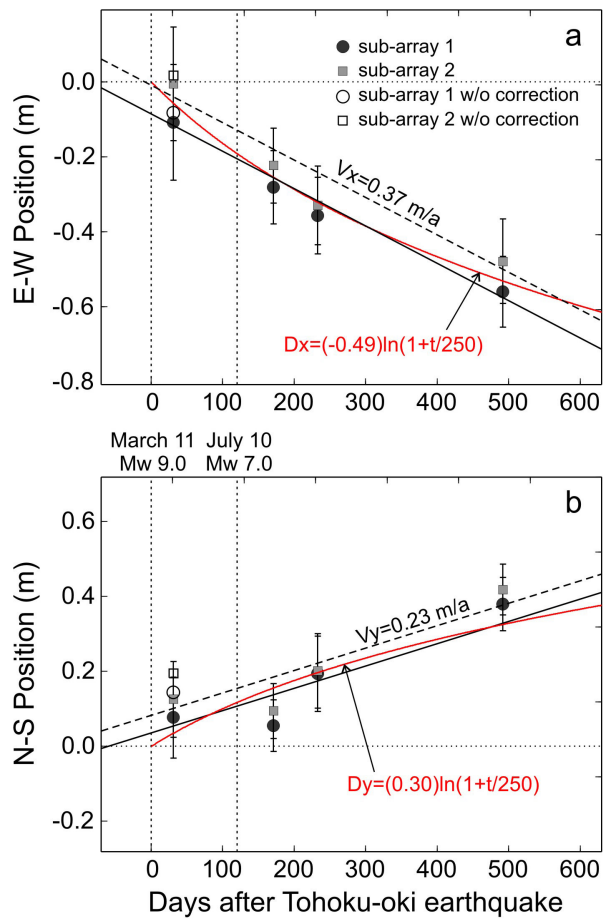




**Extended Data Figure 4 | East component of postseismic displacements of model B in Extended Data Table 1.** Otherwise the figure is the same as Fig. 3. Locations of the GPS sites are shown in Extended Data Fig. 3.



**Extended Data Figure 5** | Layout of PXPs (precision transponders) at seafloor GPS site GJT3. Grey filled circles are PXPs installed for testing purposes<sup>9</sup>, not used in this work.



**Extended Data Figure 6 | Postseismic survey results for seafloor GPS site GJT3. a, East component Dx. b, North component Dy.** Open symbols for the first measurement show array position before the effect of the  $M_w$  7.0 intraslab earthquake on 10 July 2011 was removed. Sub-array 1 includes PXP EJ12, EJ13 and EJ15, and sub-array 2 includes PXP EJ12, EJ13 and EJ22 (Extended Data Fig. 5). The straight solid and dashed lines show linear trends of survey results of sub-array 1 and sub-array 2, respectively, with resultant average velocities  $V_x$  and  $V_y$  for the east and north components, respectively. The red curves show a logarithmic function fit to the survey results.

Extended Data Table 1 | 3D model parameters

Model	Mantle wedge		Oceanic mantle		LAB layer	
	viscosity (Pa s)		viscosity (Pa s)		viscosity (Pa s)	
	$\eta_K$	$\eta_M$	$\eta_K$	$\eta_M$	$\eta_K$	$\eta_M$
A	$2.5 \times 10^{17}$	$1.8 \times 10^{18}$	$2.0 \times 10^{18}$	$1.0 \times 10^{20}$	$2.5 \times 10^{17}$	$2.5 \times 10^{17}$
B	$5.0 \times 10^{17}$	$1.0 \times 10^{19}$	$5.0 \times 10^{17}$	$1.0 \times 10^{20}$	Not applicable	

Here  $\eta_K$  and  $\eta_M$  are transient (Kelvin) and steady-state (Maxwell) viscosities, respectively. In both models, the elastic upper plate landward of the cold nose (Fig. 2c) and the lower plate are of thicknesses 25 km and 45 km, respectively, both with rigidity 48 GPa. Rigidity of the Maxwell body of the viscoelastic mantle is 64 GPa. Rigidity of the Kelvin body is 136 GPa in model A and 64 GPa in model B (the same as ref. 1). The Poisson's ratio and rock density are 0.25 and  $3,300 \text{ kg m}^{-3}$ , respectively. LAB, lithosphere–asthenosphere boundary.



# Molecular basis of adaptation to high soil boron in wheat landraces and elite cultivars

Margaret Pallotta<sup>1\*</sup>, Thorsten Schnurbusch<sup>1,2\*</sup>, Julie Hayes<sup>1</sup>, Alison Hay<sup>1</sup>, Ute Baumann<sup>1</sup>, Jeff Paull<sup>3</sup>, Peter Langridge<sup>1</sup> & Tim Sutton<sup>1</sup>

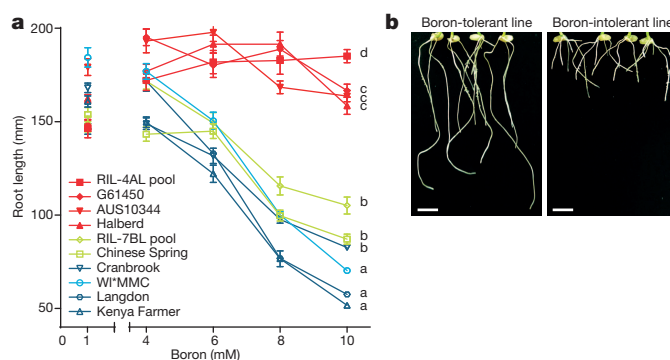
Environmental constraints severely restrict crop yields in most production environments, and expanding the use of variation will underpin future progress in breeding. In semi-arid environments boron toxicity constrains productivity, and genetic improvement is the only effective strategy for addressing the problem<sup>1</sup>. Wheat breeders have sought and used available genetic diversity from landraces to maintain yield in these environments; however, the identity of the genes at the major tolerance loci was unknown. Here we describe the identification of near-identical, root-specific boron transporter genes underlying the two major-effect quantitative trait loci for boron tolerance in wheat, *Bo1* and *Bo4* (ref. 2). We show that tolerance to a high concentration of boron is associated with multiple genomic changes including tetraploid introgression, dispersed gene duplication, and variation in gene structure and transcript level. An allelic series was identified from a panel of bread and durum wheat cultivars and landraces originating from diverse agronomic zones. Our results demonstrate that, during selection, breeders have matched functionally different boron tolerance alleles to specific environments. The characterization of boron tolerance in wheat illustrates the power of the new wheat genomic resources to define key adaptive processes that have underpinned crop improvement.

Although wheat is the most important source of calories and protein for much of the world's population, its large and complex genome has made it recalcitrant to molecular technologies. However, in comparison with model organisms, wheat has the advantages of extensive monitoring and archiving of genotypes and associated phenotypic data and the availability of unique populations adapted to specific environments and end-uses that have resulted from a long history of selective breeding. Early farmers and then modern breeders selected lines adapted to specific environments but, as with most crops, only a small proportion of the available variation in landraces and wild relatives has been effectively captured in breeding programmes<sup>3</sup>. Understanding the molecular basis for key adaptive traits would provide a powerful strategy for targeting new sources of variation. Rapidly expanding wheat genomic resources are now improving the tractability of molecular studies in wheat. Here we show the power of coupling germplasm collections with genomic resources by investigating wheat adaptation to high concentrations of boron in the soil.

In plants, boron is essential but has a narrow optimal range. Boron toxicity occurs in dry environments, often where plants are grown on alkaline soils of marine or volcanic origin, but sometimes also as a consequence of irrigation<sup>4</sup>. In bread wheat (*Triticum aestivum* L.; genomes AABBDD), boron toxicity results in decreases in root growth (Fig. 1a, b), above-ground biomass and yield. Conversely, boron deficiency is associated with high-rainfall climates and soils prone to depletion of mobile elements, resulting in poor seed set or sterility<sup>5</sup>. In a large study involving 233 trials across Australia over 12 years<sup>6</sup>, boron-tolerant genotypes had an up to 16% yield advantage over intolerant genotypes in southern wheat-growing regions where boron toxicity has been noted, but a yield disadvantage at sites in northern regions, demonstrating environment-specific adaptation. Quantitative trait loci (QTL) associated with boron tolerance

are known in wheat<sup>2,7</sup> and barley<sup>8</sup>, but the only cereal genes that have been shown to have a significant role in tolerance are from barley. In barley the 4H tolerance gene (*HvBot1*) encodes an anion-permeable transporter that is tandemly duplicated and highly expressed in the tolerant line<sup>9</sup>, and the 6H tolerance gene (*HvNIP2;1*) encodes a member of the NIP aquaporin family<sup>10</sup>. Similarly, in rice and *Arabidopsis*, genes of these two families have been implicated in responses to boron supply<sup>11</sup>. However, direct orthologues of these genes do not co-locate with *Bo1* or *Bo4*, the major boron tolerance QTL in wheat. *Bo1* is located on chromosome 7BL in the bread wheat cultivar Halberd<sup>7,12</sup> and the durum wheat (*T. turgidum* L. var. durum; genomes AABB) cultivar Lingzhi Baimong Baidamai (abbreviated as Lingzhi)<sup>13</sup>. *Bo4* is located on chromosome 4AL in the bread wheat landrace G61450 (ref. 1).

*Bot-D2a* (TaBor2, GenBank accession number EU220225; Extended Data Table 1) was previously proposed to be responsible for boron tolerance in bread wheat<sup>14</sup>, and boron transporter sequences *Bot-A4a*, *Bot-B4a* and *Bot-D4a* (*TaBOR1.2*, *TaBOR 1.3* and *TaBOR 1.1*) were described recently<sup>15</sup>. Wheat gene nomenclature guidelines are described in the Supplementary Discussion. Here we show that *Bot-D2a* maps to chromosome group 3 (Extended Data Fig. 1a), and *Bot-A4a*, *Bot-B4a* and *Bot-D4a* locate to group 5 (International Wheat Genome Sequencing Consortium website, <http://www.wheatgenome.org>) and are not associated with major loci involved in boron tolerance. We previously reported a fine map of the 7BL region containing *Bo1* in the bread wheat doubled haploid population Cranbrook (intolerant) × Halberd<sup>16</sup>. Here we genotyped 1,700 individuals with markers developed from genes in syntenic regions on *Brachypodium* supercontig 1 and rice chromosome 6 to identify



**Figure 1 | Effect of *Bo1* allele type on root length at high boron concentrations.** **a**, Root lengths of seedlings of representative genotypes grown at a range of concentrations of boron in hydroponics ( $n = 16$ ; means  $\pm$  s.e.m.). Letters denote significant ( $P < 0.01$ ) differences between genotypes for linear regression analysis across the range 4–10 mM boron. RIL-4AL pool and RIL-7BL pool are each pools of eight G61450 × Kenya Farmer recombinant inbred lines. **b**, Roots of tolerant (Halberd) and intolerant (Cranbrook) wheat genotypes grown for 10 days in hydroponics containing 10 mM boron. Scale bars, 10 mm.

<sup>1</sup>Australian Centre for Plant Functional Genomics, School of Agriculture, Food and Wine, University of Adelaide, Waite Campus, Urrbrae, South Australia 5064, Australia. <sup>2</sup>Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Genebank Department, Corrensstrasse 3, D-06466 Gatersleben, Germany. <sup>3</sup>School of Agriculture, Food and Wine, University of Adelaide, Waite Campus, Urrbrae, South Australia 5064, Australia.

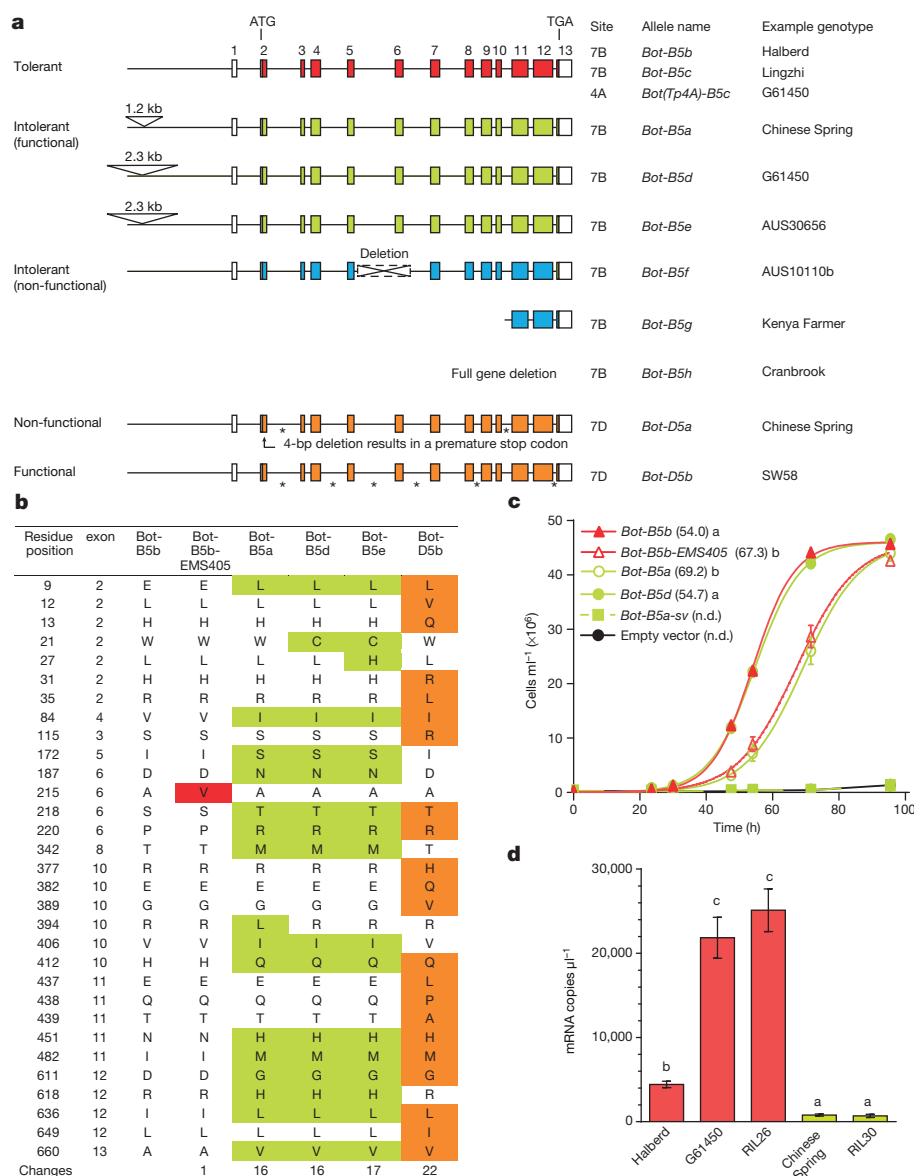
\*These authors contributed equally to this work.

153 lines recombinant between *barc32* and AWW355. For each F<sub>2</sub> recombinant, boron tolerance of F<sub>3</sub> progeny families was assessed by root growth in high boron hydroponics. This reduced the target interval containing the tolerance gene to 0.06 centimorgans (Extended Data Fig. 2a). Gene annotation in the syntenic interval from related grass species did not reveal candidate genes for *Bo1*.

In parallel, we screened BACs derived from the orthologous interval in *Aegilops tauschii* (genome DD)<sup>17</sup> with a probe derived from *HvBot1* (ref. 9). A positive clone (H1148P11) was used to identify the sequence of a D-genome *HvBot1*-like gene. A genomic DNA fragment (AWW461) of this gene co-segregated with *Bo1*. Sequencing of the gene from Halberd showed that the 7BL tolerance locus contains an undescribed boron transporter-like gene with 80% open reading frame (ORF) similarity to *HvBot1* and *Bot-D2a*. The Halberd gene (*Bot-B5b*; Fig. 2a) contains 12 introns and encodes a predicted membrane protein of 660 amino-acid residues (GenBank accession number KF148625). We confirmed membrane localization *in planta* by confocal imaging of onion epidermal cells transiently expressing the construct 35S:*Bot-B5b*:GFP (Extended Data Fig. 3a), and heterologous expression in *Saccharomyces cerevisiae* indicated that *Bot-B5b* is able to function as a boron transporter (Fig. 2c). Using plant boron transporter ORF sequences we constructed the maximum-likelihood phylogeny of this gene family, verifying that *Bot-B5*

genes have no direct orthologous sequences in barley (Supplementary Discussion) or in the sequenced reference genotypes of *Brachypodium distachyon*, *Oryza sativa* (ssp. *japonica*) or *Sorghum bicolor* (Extended Data Fig. 4a, b).

Gene similarity between bread wheat *Bo1* and *Bo4* was investigated by mapping AWW461 in recombinant inbred lines (RILs) derived from a cross G61450 × Kenya Farmer, in which *Bo4* was previously linked to *XksuG10-4A* (ref. 18). We found two copies, one explaining 79% of total trait variation for absolute root length under high boron in the 4AL interval *Xabg390-4A-XksuG10-4A*, the other locating to the *Bo1* locus on 7BL with no significant marker-trait association (Extended Data Fig. 2b). Investigation of the extent of localized similarity around the Halberd 7BL and G61450 4AL genes showed no evidence of chromosomal translocation, indicating that *Bo4* represents a dispersed duplication<sup>19</sup> of the 7BL gene. Similarly, in durum wheat we mapped a *Bot-B5b*-derived marker (AWW555-SacI) in F<sub>2</sub> plants of Jandaroi (intolerant) × AUS14740 (tolerant) and found co-segregation with AWW5L7, a 7BL-specific marker tightly linked to *Bo1* (ref. 20), supporting orthology of the bread and durum wheat 7BL loci. Our findings implicate boron transporters in boron tolerance at three major tolerance loci: *Bo1* in bread and durum wheat, and *Bo4* in bread wheat. Across 9.4 kilobases (kb) of genomic sequence, G61450 4AL and Lingzhi 7BL genes are identical and differ from Halberd 7BL by a single non-synonymous



**Figure 2 | Variation in *Bot-B5/D5* alleles.** Alleles are grouped by colour. **a**, Schematic diagrams showing the gene structure of *Bot-B5/D5* alleles. In *Bot-D5a* and *Bot-B5b*, asterisks indicate undetermined intron sizes. **b**, Residue differences between functional *Bot-B5/D5* alleles. Changes relative to *Bot-B5b* are shaded. **c**, Growth of yeast expressing *Bot-B5* variants and control sequences, in liquid culture supplemented with 15 mM boron ( $n = 5$ ; means  $\pm$  s.e.m.). Hours to half-saturation are given after each yeast strain (n.d., not determined), and letters denote statistically distinct groups ( $P < 0.01$ ). No significant difference was observed for yeast expressing *Bot-B5* variants at low boron concentrations. **d**, *Bot-B5* transcript level ( $n = 4$ ; means  $\pm$  s.e.m.) in roots exposed to 2 mM boron for 22 h: Halberd, G61450 (*Bot(Tp4A)-B5c* + *Bot-B5d*), RIL26 (*Bot(Tp4A)-B5c* + *Bot-B5g*), Chinese Spring and RIL30 (*Bot-B5d*). Letters denote significant ( $P < 0.01$ ) differences between genotypes for  $\log_{10}$ -transformed data.

nucleotide. High sequence identity together with pedigree information (Genetic Resources Information System for Wheat and Triticale, <http://www.wheatpedigree.net/>) implicates tetraploid lines as sources of *Bo1* and *Bo4* in hexaploid bread wheat (Supplementary Discussion). Analysis of neighbouring genes on 7BL in current bread wheat cultivars indicates conservation of the tetraploid-derived segment, presenting a barrier to recombination and the exploitation of agronomically important loci linked to *Bo1*, such as resistance to late-maturity  $\alpha$ -amylase<sup>21</sup>.

A population mutagenized with ethyl methanesulphonate (EMS) was developed in Halberd. At high boron, two independent mutants with intermediate root growth were identified (Extended Data Fig. 5a). Mutant EMS405 contained a substitution in exon 6 of *Bot-B5b* (Ala 215→Val), which reduced yeast growth under high boron in a complementation assay (Fig. 2c). Mutant EMS388 contained a single nucleotide substitution (G in wild type, A in mutant) within the *Bot-B5b* gene promoter sequence, 215 base pairs (bp) upstream of the predicted messenger RNA start site. No difference in expression of the *Bot-B5b* transcript was observed between EMS388 and Halberd. Both mutants co-segregated with reduced root growth under high boron (Extended Data Fig. 5b).

Bread and durum wheat *Bot-B5* alleles broadly fell into three groups. The first comprised all boron-tolerant cultivars containing the Halberd (*Bot-B5b*), G61450 (*Bot(Tp4A)-B5c*) or Lingzhi (*Bot-B5c*) alleles. Intolerant lines fell into two further allele groups: those related to the gene in the reference genotype Chinese Spring (*Bot-B5a*), and those in which the gene was partly or wholly deleted (Fig. 2a). In Cranbrook (hexaploid) and Langdon (tetraploid), which contain the null allele *Bot(Df)-B5h* (Extended Data Fig. 1b), the genomic deletion is estimated at more than 22 kb and includes the complete *Bot-B5* gene. Chinese Spring-group alleles (*Bot-B5a*, *Bot-B5d* and *Bot-B5e*) have 98% ORF sequence identity to *Bot-B5b* and are characterized by the insertion of repetitive sequences in the promoter region 2,240 bp upstream of the mRNA start codon site, in addition to protein sequence polymorphism compared with Halberd *Bot-B5b* (Fig. 2a, b). Protein function comparison in yeast under high boron showed reduced function of *Bot-B5a* compared with both *Bot-B5b* and *Bot-B5d* (Fig. 2c). *Bot-B5a* and *Bot-B5d* differ by only two residues, suggesting key roles of one or both residues. Alleles in the Halberd and Chinese Spring groups both showed root-specific expression that is responsive to high boron (Extended Data Fig. 3b, c), but they differed in transcript level (Fig. 2d), consistent with variation in observed root length phenotypes under high boron (Fig. 1a). The low level of expression found for the Chinese Spring group may have been due to the insertions in the promoter sequence. High expression of *Bot-B5* in G61450 was derived from the 4AL allele *Bot(Tp4A)-B5c*, illustrated by comparison between G61450 × Kenya Farmer-derived lines RIL26 and RIL30 (Fig. 2d). *Bot-B5c* in durum wheat and *Bot(Tp4A)-B5c* in G61450 showed similar expression levels, consistent with sequence identity and recent transposition of a functional gene.

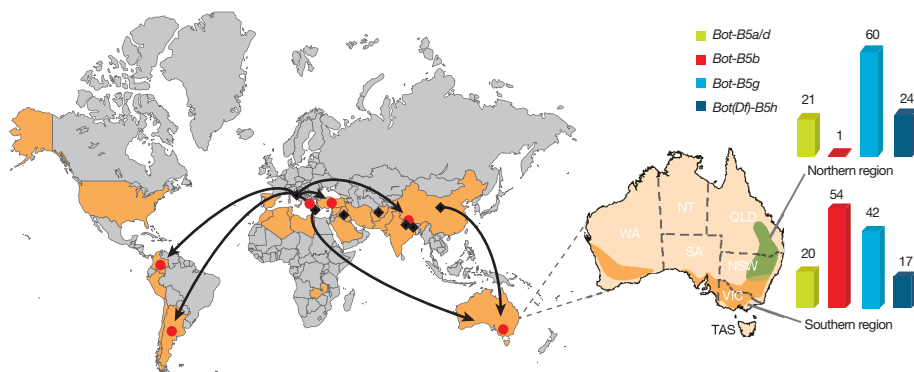
No homoeologous 7A sequences have been found in bread wheat, in durum wheat or in A genome progenitor species, supporting the observation<sup>20</sup>

that boron tolerance is absent among *T. monococcum* and *T. urartu* accessions, and implying an *Aegilops* source for this boron transporter in cultivated wheat. In Chinese Spring (see <http://www.wheatgenome.org/>), we found a D genome sequence (*Bot-D5a*) low in expression that contained a 4-bp frame-shift mutation encoding a truncated, non-functional protein of 56 residues (Fig. 2a, b). We originally identified *Bot-B5b* through the bridging genotype *Ae. tauschii* accession AL8/78, so examined the synthetic hexaploid wheat SW58 derived from Langdon (intolerant) × AL8/78. We identified a transcript (*Bot-D5b*) more highly expressed than *Bot-D5a* with 97% protein identity to Halberd *Bot-B5b* (Fig. 2a, b). In a yeast complementation assay, *Bot-D5b* was functional but had lower efficacy than *Bot-B5b*, supporting plant root length data that demonstrate less tolerance in SW58 than Halberd, but greater tolerance than Langdon (Extended Data Fig. 6a, b).

In total we identified 12 sequence variants for the *Bot-B5/D5* genes, and we can account for boron tolerance phenotype on the basis of gene sequence: boron intolerance results from a loss of function through complete or partial deletion, or frame shift mutation (Cranbrook, AUS10110b, Kenya Farmer and Chinese Spring *Bot-D5a*), partial loss of function or reduced effect from induced mutation (EMS405 and EMS388), and decreased transcript level (Chinese Spring *Bot-B5a*, G61450 *Bot-B5d* and AUS30656 *Bot-B5e*). Three highly conserved but distinct natural variants (Halberd, Lingzhi and G61450 *Bot(Tp4A)-B5c*) are fully functional.

The adaptive advantages of different boron tolerance alleles was demonstrated by genotyping a set of 85 released cultivars and 153 advanced breeding lines, revealing a biased deployment of *Bot-B5* alleles between southern and northern Australian wheat-growing regions (Fig. 3): tolerance alleles predominated in southern regions and were absent in lines targeted to northern regions. Despite relatively early introduction into Australian breeding programmes, the highly expressed *Bot(Tp4A)-B5c* allele was not detected in lines from either region. Furthermore, genotyping of boron-tolerant lines from diverse locations outside Australia failed to identify lines carrying both *Bot-B5b* and *Bot(Tp4A)-B5c*, consistent with a penalty associated with the presence of strong tolerance alleles in low-boron environments.

The challenge in expanding the variation available to breeders depends on the identification and subsequent deployment of novel variation. Selection of wheat lines adapted to different production environments has been occurring since wheat was domesticated about 10,000 years ago; in selecting for performance, early farmers around the Mediterranean, through the Middle East and into northern India, Afghanistan and China developed wheat landraces with varying levels of tolerance to high concentrations of boron in the soil (Fig. 3). A similar process was occurring in barley, but our data reveal that boron tolerance in barley and wheat arose through the divergent evolution of paralogous genes (Extended Data Fig. 4a, b). In wheat the generation of comparatively broad allelic variation provided adaptation to agro-geographically diverse regions. Plant breeders in the early twentieth century recognized the value of landraces as sources of useful variation and exploited locally adapted lines, leading to the emergence



**Figure 3 | Wheat *Bot-B5* allele origin and dispersion, and Australian distribution pattern.**

Black diamonds indicate predicted sources of the tolerance alleles *Bot-B5b*, *Bot(Tp4A)-B5c* and *Bot-B5c*, with proposed dispersion shown by black arrows. Red circles show countries where modern cultivars carrying tolerance alleles have been found. Orange shading indicates countries where boron toxicity has been identified<sup>4,25,26</sup>. Within the enlarged map of Australia, coloured areas broadly depict southern and northern wheat-growing regions. Numbers above the graph bars show the number of lines for each allele type. The single northern-bred cultivar identified to contain *Bot-B5b* was bred for adaptation to southern environments. The *Bot(Tp4A)-B5c* allele was not detected in screened Australian germplasm.



of boron-tolerant varieties in the Mediterranean. Wheat is a relatively recent introduction into Australia and the Americas, with early varieties based on lines adapted to temperate, usually northern European, environments. Problems of adaptation in these new environments led breeders back to landraces, where diversity was sought in accessions derived from regions with perceived environmental similarities (Fig. 3). In this study we have shown that for both bread and durum wheat this empirical approach was successful and resulted in the release of boron-tolerant cultivars in Australia and South America. We have also shown that there has been strong selection for or against functionally different *Bot-B5* alleles in contrasting environments, which implies that matching boron tolerance alleles to the level of soil boron is critical in achieving maximum yield potential.

## METHODS SUMMARY

**Phenotyping for boron toxicity response.** The length of the longest root of 8–14-day-old seedlings grown either in aerated hydroponics or on moistened filter paper was measured as described<sup>18,20</sup> except that in the hydroponics assays the low-boron treatment was 0.015 mM boron.

**Molecular biology.** Genetic manipulations followed established protocols. Sequence data were obtained from 454 sequencing of Chinese Spring BAC clones 112M01 and 451O08 (ref. 22), from Sanger sequencing of genomic and cDNA fragments and from databases of wheat genomic sequences<sup>23</sup>. Gene expression analysis of allelic variants used four biological replicates comprising root tissue from seedlings grown hydroponically for 16 days in 0.05 mM boron followed by 22 h in 2 mM boron.

**Functional assessment in yeast.** Yeast strain INVSc2 was used in all experiments. Growth experiments in liquid medium and on solid medium were as described previously<sup>9</sup>. Data from time-course growth assays at low and high boron concentrations were plotted and fitted to Boltzmann sigmoidal functions by using nonlinear modelling (GraphPad Prism 6) to calculate times to half-saturation for five replicates. **Transient expression in onion epidermal cells.** The *Bot-B5b* coding sequence was cloned into vector pMDC83 to generate a carboxy terminus construct, 35S:*Bot-B5b*:GFP, and introduced into onion epidermis by bombardment. Cells were visualized by confocal image analysis before and after plasmolysis.

**Phylogeny.** Complete ORF sequences were aligned using MUSCLE, and the maximum likelihood based on the Kimura 2 parameter model was calculated. Bootstrap values were generated from 1,000 replicates.

**Genetic mapping.** Mapping and QTL analyses were performed with MapManager QTX version 0.30 (ref. 24).

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 October 2013; accepted 28 May 2014.

Published online 2 July 2014.

1. Paull, J. G., Nable, R. O. & Rathjen, A. J. Physiological and genetic control of the tolerance of wheat to high concentrations of boron and implications for plant breeding. *Plant Soil* **146**, 251–260 (1992).
2. Paull, J. G., Rathjen, A. J. & Cartwright, B. Major gene control of tolerance of bread wheat (*Triticum aestivum* L.) to high concentrations of soil boron. *Euphytica* **55**, 217–228 (1991).
3. McCouch, S. *et al.* Agriculture: feeding the future. *Nature* **499**, 23–24 (2013).
4. Nable, R. O., Banuelos, G. S. & Paull, J. G. Boron toxicity. *Plant Soil* **193**, 181–198 (1997).
5. Rerkasem, B. & Jamjod, S. Boron deficiency induced male sterility in wheat (*Triticum aestivum* L.) and implications for plant breeding. *Euphytica* **96**, 257–262 (1997).
6. McDonald, G. K., Taylor, J. D., Verbyla, A. & Kuchel, H. Assessing the importance of subsoil constraints to yield of wheat and its implications for yield improvement. *Crop Pasture Sci.* **63**, 1043–1065 (2012).

7. Jefferies, S. P. *et al.* Mapping and validation of chromosome regions conferring boron toxicity tolerance in wheat (*Triticum aestivum*). *Theor. Appl. Genet.* **101**, 767–777 (2000).
8. Jefferies, S. P. *et al.* Mapping of chromosome regions conferring boron toxicity tolerance in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **98**, 1293–1303 (1999).
9. Sutton, T. *et al.* Boron-toxicity tolerance in barley arising from efflux transporter amplification. *Science* **318**, 1446–1449 (2007).
10. Schnurbusch, T. *et al.* Boron toxicity tolerance in barley through reduced expression of the multifunctional aquaporin HvNIP2;1. *Plant Physiol.* **153**, 1706–1715 (2010).
11. Miwa, K. & Fujiwara, T. Boron transport in plants: co-ordinated regulation of transporters. *Ann. Bot. (Lond.)* **105**, 1103–1108 (2010).
12. Chantachume, Y. *Genetic Studies on the Tolerance of Wheat to High Concentrations of Boron*. DPhil thesis, Univ. Adelaide (1995).
13. Jamjod, S. *Genetics of Boron Tolerance in Durum Wheat*. DPhil thesis, Univ. Adelaide (1996).
14. Reid, R. Identification of boron transporter genes likely to be responsible for tolerance to boron toxicity in wheat and barley. *Plant Cell Physiol.* **48**, 1673–1678 (2007).
15. Leangthitikanachana, S. *et al.* Differential expression of three *BOR1* genes corresponding to different genomes in response to boron conditions in hexaploid wheat (*Triticum aestivum* L.). *Plant Cell Physiol.* **54**, 1056–1063 (2013).
16. Schnurbusch, T. *et al.* Fine mapping and targeted SNP survey using rice–wheat gene colinearity in the region of the *Bo1* boron toxicity tolerance locus of bread wheat. *Theor. Appl. Genet.* **115**, 451–461 (2007).
17. Luo, M.-C. *et al.* A 4-gigabase physical map unlocks the structure and evolution of the complex genome of *Aegilops tauschii*, the wheat D-genome progenitor. *Proc. Natl Acad. Sci. USA* **110**, 7940–7945 (2013).
18. Paull, J. G., Rathjen, A. J., Langridge, P. L. & McIntosh, R. A. in *Proceedings of the 8th International Wheat Genetics Symposium, Beijing, China, 20–25 July 1993* (eds Li, Z. S. & Xin, Z. Y.) 1065–1069 (China Agricultural Science Press, 1993).
19. Akhunov, E. D., Akhunova, A. R. & Dvorak, J. Mechanisms and rates of birth and death of dispersed duplicated genes during the evolution of a multigene family in diploid and tetraploid wheats. *Mol. Biol. Evol.* **24**, 539–550 (2007).
20. Schnurbusch, T., Langridge, P. & Sutton, T. The *Bo1*-specific PCR marker AWW5L7 is predictive of boron tolerance status in a range of exotic durum and bread wheats. *Genome* **51**, 963–971 (2008).
21. Mares, D. & Mrva, K. Late-maturity  $\alpha$ -amylase: low falling number in wheat in the absence of preharvest sprouting. *J. Cereal Sci.* **47**, 6–17 (2008).
22. Allouis, S. *et al.* Construction and characterisation of a hexaploid wheat (*Triticum aestivum* L.) BAC library from the reference germplasm ‘Chinese Spring’. *Cereal Res. Commun.* **31**, 331–338 (2003).
23. Brechley, R. *et al.* Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature* **491**, 705–710 (2012).
24. Manly, K. F., Cudmore, R. H. Jr & Meer, J. M. Map Manager QTX, cross-platform software for genetic mapping. *Mamm. Genome* **12**, 930–932 (2001).
25. Mathews, K. L. *et al.* Indirect selection using reference and probe genotype performance in multi-environment trials. *Crop Pasture Sci.* **62**, 313–327 (2011).
26. Yau, S. K. & Ryan, J. Boron toxicity tolerance in crops: a viable alternative to soil amelioration. *Crop Sci.* **48**, 854–865 (2008).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank N. Collins, J. Dvorjak, H. Kuchel, D. Mares, S. Wu, the John Innes Centre, the Biotechnology and Biological Sciences Research Council, the Institut National de la Recherche Agronomique and the International Wheat Genome Sequencing Consortium for resources, and J. Tiong, T. Oz and A. Pohlen for assistance. The authors are supported by grants from the Australian Research Council, the Grains Research and Development Corporation and the South Australian Government.

**Author Contributions** M.P., T.Sc., J.H., J.P., P.L. and T.S. designed experiments. M.P., T.Sc., J.H. and A.H. performed experiments. M.P., T.Sc., J.H., A.H., U.B., J.P. and T.S. analysed data. M.P., T.S., J.H. and P.L. wrote the manuscript. T.Sc., U.B. and J.P. commented on the manuscript.

**Author Information** Sequence data are deposited with NCBI GenBank under accession numbers KF148623–KF148633 and GF112200–GF112209. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.S. ([tim.sutton@acpfg.com.au](mailto:tim.sutton@acpfg.com.au)).



## METHODS

**Plant and DNA materials.** Bread wheat lines used for phenotyping assays and whole-genome sequencing were Cranbrook, Chinese Spring, G61450, Halberd, Kenya Farmer, WI\*MMC, and lines from a G61450 × Kenya Farmer recombinant inbred line (RIL) mapping population of 90 lines<sup>18</sup> (RIL26, RIL30, and tolerant and intolerant G61450 × Kenya Farmer RIL pools). RIL26 and RIL-4AL pool (RIL6, RIL7, RIL9, RIL26, RIL62, RIL66, RIL87 and RIL91) contain a G61450 allele at *Bo4* on chromosome 4AL (functional allele *Bot(Tp4A)-B5c*) and a Kenya Farmer allele at *Bo1* on chromosome 7BL (truncated allele *Bot-B5g*). RIL30 and RIL-7BL pool (RIL16, RIL21, RIL27, RIL30, RIL44, RIL56, RIL93 and RIL97) have a Kenya Farmer allele at *Bo4* on 4AL (null) and a G61450 allele at *Bo1* on 7BL (functional allele *Bot-B5d*). Durum wheat lines used for phenotyping assays and whole-genome sequencing were Langdon and the landraces AUS10110 (Uttar Pradesh, India), AUS10344 (*Triticum durum* Desf. var. *niloticum*, Iraq), AUS14010 (Lingzhi Baimong Baidamai, China) and AUS14740 (Afghanistan). AUS10110, AUS10344 and AUS14010 were previously identified as boron-tolerant in a screen<sup>13</sup> of 300 genotypes from North Africa, Asia, Australia, Italy and the International Maize and Wheat Improvement Center (CIMMYT). In our study, AUS10110 was identified as segregating for boron tolerance and found to be heterogeneous at the 7B *Bot-B5* locus. Selections of AUS10110 were made based on allele composition, with AUS10110a containing the functional allele *Bot-B5c* and AUS10110b containing the truncated and non-functional allele *Bot-B5f*. Synthetic wheats used for phenotyping assays, genetic analysis and whole-genome sequencing were SW58 (Langdon × *Ae. tauschii* AL8/78) and AUS30656 (LCK59.61/*Ae. tauschii*). SW58 was supplied by S. Wu.

Cultivars, breeders' lines or DNA samples for marker screening were obtained from various Australian wheat breeding programmes, D. Mares and the Australian Winter Cereals Collection (AWCC). DNA of Turkish cultivars was supplied by T. Oz. DNA of 112 F<sub>2</sub> lines from the tetraploid population Jandaroi × AUS14740 was supplied by N. Shamaya.

Halberd-EMS mutagenized lines were generated by treatment of wheat cv. Halberd seed with 0.45% (v/v) ethyl methanesulphonate (EMS) for 16 h. Putative mutant EMS-Halberd M<sub>3</sub> seedlings were identified by phenotyping about 15–20 seeds per M<sub>2</sub> family in high-boron hydroponics (10 mM boron) in a greenhouse for 10–14 days. Boron-intolerant families were identified on the basis of short root length and the appearance of the first leaf (tip necrosis, yellow wilting). Individual putative mutant plants were selected and transplanted into hydroponics without boron for recovery. After approximately 2 weeks of recovery hydroponics, survivors were transplanted into pots of soil for seed multiplication. Further phenotype validation undertaken in M<sub>4</sub> and M<sub>5</sub> generations and full gene sequencing of M<sub>4</sub> pools resulted in the identification of two *Bot-B5b* mutant families, EMS388 and EMS405.

**Hydroponic phenotyping for boron toxicity response.** Seeds for phenotyping assays, with the exception of the G61450 × Kenya Farmer RIL population, were germinated on filter paper and grown hydroponically in a greenhouse or controlled-environment chamber (20 °C day/15 °C night, 12-h day) for 1 day in a low-boron minimal nutrient base solution containing 0.5 mM Ca(NO<sub>3</sub>)<sub>2</sub>, 2.5 μM ZnSO<sub>4</sub> and 0.015 mM boron as H<sub>3</sub>BO<sub>3</sub> and then for 8–14 days, with aeration, in either fresh base solution (low-boron treatment) or base solution plus additional boron (4–10 mM boron) as H<sub>3</sub>BO<sub>3</sub>. Solutions were replaced once or twice, depending on the length of treatment. Seedlings showing fungal infection around the seed were discarded. Absolute length of the longest root (RL) was measured for 13–16 seedlings per line at each treatment, and average RL and s.e.m. were calculated. Previous studies have shown that RL under high boron is highly correlated with boron tolerance<sup>12,20</sup>. Phenotyping of the G61450 × Kenya Farmer RIL population was performed similarly, except that the seedlings were grown in 9.3 mM boron (100 mg kg<sup>-1</sup> boron) on filter paper as described previously<sup>27</sup>.

Plants for quantitative RT-PCR (qRT-PCR) and northern analysis were grown in two experiments (experiments 1 and 2). In both experiments, plants were grown for 17 days in a full nutrient base hydroponics solution containing 5 mM NH<sub>4</sub>NO<sub>3</sub>, 5 mM KNO<sub>3</sub>, 2 mM Ca(NO<sub>3</sub>)<sub>2</sub>, 2 mM MgSO<sub>4</sub>, 0.1 mM KH<sub>2</sub>PO<sub>4</sub>, 0.05 mM NaFe(III) EDTA, 0.05 mM H<sub>3</sub>BO<sub>3</sub>, 5 μM MnCl<sub>2</sub>, 10 μM ZnSO<sub>4</sub>, 0.5 μM CuSO<sub>4</sub> and 0.1 μM Na<sub>2</sub>MoO<sub>4</sub> with aeration, in a controlled-environment growth room at 22 °C (day)/16 °C (night) with a 14-h photoperiod. Solutions were replaced every 3–4 days during the experiment. Seeds were germinated on filter paper, and seedlings with shoots of 2–3 cm were transplanted to hydroponics. The experiment design was four biological replicates per genotype for each of three treatments, and seedlings within a treatment were arranged in a modified Latin-square pattern. In experiment 1, each biological replicate comprised a pool of two seedlings to reduce sampling error further. In experiment 2, each biological replicate comprised a single seedling. Treatments were low boron for 17 days, 22 h at 2 mM boron in full nutrient solution applied at day 16, and 7 days at 2 mM boron in full nutrient solution applied at day 10. All treatments were harvested on day 17.

**Mapping and QTL analysis.** Marker linkage in the Jandaroi × AUS14740 F<sub>2</sub> population, and marker linkage and single marker regression QTL analyses in the G61450 ×

Kenya Farmer RIL population, were all performed with MapManager QTX version 0.30 (ref. 24). Genetic map images were generated with MapChart 2.2 software<sup>28</sup>.

**Nucleic acid extraction, Southern and northern analysis, rapid amplification of cDNA ends (RACE), cDNA synthesis and quantitative real-time PCR.** Genomic DNA was extracted using either of two standard methods<sup>29,30</sup>; phenol/chloroform-extracted DNA was used for Southern analysis. Total RNA was extracted from roots of hydroponically grown plants with TRIzol (Invitrogen) followed by ISOLATE plant RNA spin column purification (Bioline). We synthesized first-strand cDNA using Superscript III reverse transcriptase (Invitrogen) and used it as the template to amplify *Bot-B5* transcripts. qRT-PCR assays were performed with methods described previously<sup>31</sup>. SMART RACE (Clontech) cDNA synthesis was used to obtain cDNA for determining 5' and 3' mRNA sequences of Halberd and Chinese Spring *Bot-B5* transcripts. Southern and northern analysis using <sup>32</sup>P-labelled probes was performed with standard methods. Final washing after probe hybridization for both Southern and northern membranes was in 0.5 × SSC, 0.1% SDS solution for 20 min at 65 °C.

For Southern analysis to locate *Bot-D2a* in wheat (Extended Data Fig. 1a), we digested genomic DNA from Chinese Spring nullisomic-tetrasomic (CS N-T) chromosome substitution lines with DraI and hybridized with the <sup>32</sup>P-labelled probe AWW469, a 261-bp cDNA fragment of *Bot-D2a* amplified from Cranbrook, which does not cross-hybridize to *Bot-B5* or *Bot-D5* at high stringency. We included a genotype of the D-genome species *Ae. tauschii* to assist in interpretation. For Southern analysis to demonstrate both the absence of 7B *Bot-B5* sequences in Cranbrook and Langdon, and the absence of homoeologous 7A sequences in bread and tetraploid wheat (Extended Data Fig. 1b), we used the bread wheat cultivars Cranbrook, Halberd, Chinese Spring and CS N7B-T7A, and the reference durum wheat cultivar Langdon. Genomic DNA was digested with HindIII and hybridized with the <sup>32</sup>P-labelled probe AWW471, a 357-bp genomic DNA fragment of *Bot-D5b* that hybridizes to *Bot-B5* at high stringency.

For northern analysis of *Bot-B5* transcript induction under high-boron conditions (Extended Data Fig. 3c) we used root tissue from Halberd grown in two independent experiments (experiment 1 and experiment 2, described in detail above) and G61450 seedlings grown in one of the experiments (experiment 1). To increase replication in experiment 2, where each biological replicate comprised a single plant, two sets of Halberd lines were sampled. Hybridization with a 268-bp cDNA probe (AWW548) derived from Halberd, comprising 65 bp of coding sequence and 203 bp of 3' UTR, was used to detect *Bot-B5* transcripts.

Semi-quantitative RT-PCR was performed on a Chinese Spring developmental tissue series<sup>32</sup> using the *Bot-B5*-specific marker qRT-PCR-*Bot-B5* (36 cycles) and a wheat glyceraldehyde-3-phosphate dehydrogenase (GAPDH) marker (28 cycles). PCR products were separated on 1.5% agarose in 1 × Tris-acetate-EDTA buffer and revealed with ethidium bromide by standard methods.

**Allele diversity and markers for *Bot-B5* and *Bot-D5*.** Sequence data for allelic variants of the *Bot-B5* and *Bot-D5* genes were derived from fragments amplified from both genomic DNA and cDNA templates produced from individual genotypes. PCR was generally performed using Immolase DNA polymerase (Bioline), with 36 cycles of amplification. PCR products were purified using NucleoSpin II (Macherey-Nagel) or ISOLATE (Bioline) kits. Sanger sequencing was performed using BigDye V3.1 (ABI). Sequences were aligned using Contig Express software (Vector NTI Advance 11.0, Invitrogen). Gene sequences are available under the accession numbers listed in Extended Data Table 1.

Sequencing of Chinese Spring 7B BAC clones 112M01 and 451O08 (ref. 22), in addition to accessing a database of Chinese Spring genomic sequences<sup>33</sup>, yielded a 27.3-kb contig containing *Bot-B5*. Contig authenticity was verified by PCR from a Chinese Spring genomic DNA template of fragments spanning joined sequence blocks. We confirmed sequence accuracy for a 9.4-kb genomic region spanning *Bot-B5a* by Sanger sequencing of PCR products amplified from Chinese Spring and a Chinese Spring nullisomic 7D-tetrasomic 7B template. In the same way we generated a corresponding sequence across the 9.4-kb genomic region for selected other lines. For each of these lines a transcript sequence was also obtained to verify coding region sequences and to assess variation in splice form. In all lines containing full-length *Bot-B5* alleles, a region of intron 4 contained a string of G residues about 20 bp in length that prevented through-sequencing. Sizing of fragments on 3% agarose gels, in combination with sequence data from both strands, was consistent with sequences containing only the G-string. Sequences for *Bot-D5a* and *Bot-D5b* alleles were obtained from root cDNA of Chinese Spring and SW58, respectively and were verified against genomic sequence databases and Sanger sequencing of genomic fragments.

A suite of allele-discriminating PCR markers were developed for determining the presence of *Bot-B5* alleles *a–h*, with the exception of *Bot-B5e*. *Bot-B5e* was detected in a Mexican-derived synthetic wheat line and was therefore considered unlikely to be a common allele either globally in bread wheat or in Australian tetraploid germplasm. Primer sequences are provided in Supplementary Table 1. The dominant marker AWW525, which detects only *Bot-B5b*, *Bot(Tp4A)-B5c* and *Bot-B5c*,

is based on a 24-bp duplication in the 3' UTR region, 206 bp from the TGA stop codon, and is not found in other alleles. To further distinguish *Bot-B5b* from *Bot(Tp4A)-B5c* and *Bot-B5c*, a CAPS marker, AWW532-HpyAV, based on the single exon-3 single nucleotide polymorphism was used. Co-dominant marker AWW600, based on promoter sequence differences, was used to distinguish *Bot-B5a* from *Bot-B5d/Bot-B5e* alleles. Lines having other allele types yield no AWW600 product. Lines carrying the truncated allele *Bot-B5g* or the null allele *Bot(Df)-B5h* were identified in a two-step process. The first step used the dominant marker AWW555 to identify null types, because it yields a product for all alleles except *Bot(Df)-B5h*. The second step used the dominant marker AWW516, which is located in the region of *Bot-B5* deleted in the *Bot-B5g* allele, spanning intron 4, and yielding products of different sizes from 7B and 7D genomes. Lines carrying *Bot-B5g* or *Bot(Df)-B5h* alleles yield only a single product, derived from the D-genome, whereas lines carrying full-length *Bot-B5* alleles yield two products. The use of AWW516 in conjunction with AWW555 overcame the possibility of an incorrect assignment of a null allele in the instance of a failed PCR reaction. Allele data were determined for 153 Australian advanced breeding lines, 85 Australian cultivars and 54 non-Australian hexaploid and tetraploid lines (Fig. 3).

**Functional assessment of *Bot-B5* and *Bot-D5* alleles in yeast.** Full-length coding sequences of each of *Bot-B5a*, *Bot-B5b*, *Bot-B5d*, *Bot-B5b-EMS405* and *Bot-D5b* were cloned in the Gateway entry vector pCR8 (Invitrogen). Inserts were confirmed by sequencing and transferred to a Gateway-enabled destination vector for yeast expression, pYES-DEST52 (Invitrogen).

Yeast (*Saccharomyces cerevisiae* strain INVSc2; Invitrogen) were transformed using a standard lithium acetate method<sup>34</sup>. Growth experiments on solid medium were conducted as described previously<sup>9</sup> to compare the boron tolerance of *Bot-B5b*-expressing clones and *Bot-D5b*-expressing clones with each other and with that of yeast transformed with a truncated non-functional version of the Chinese Spring *Bot-B5a* allele (*Bot-B5a-sv*). Boron tolerance of *Bot-B5*-expressing clones was quantified by culturing yeast in minimal liquid medium containing 2% galactose as a source of carbon, both at low boron and with an additional 15 mM H<sub>3</sub>BO<sub>3</sub>. Growth was recorded by removing aliquots of cell suspensions at intervals and measuring the attenuation ( $D_{600}$ ) with a spectrophotometer.

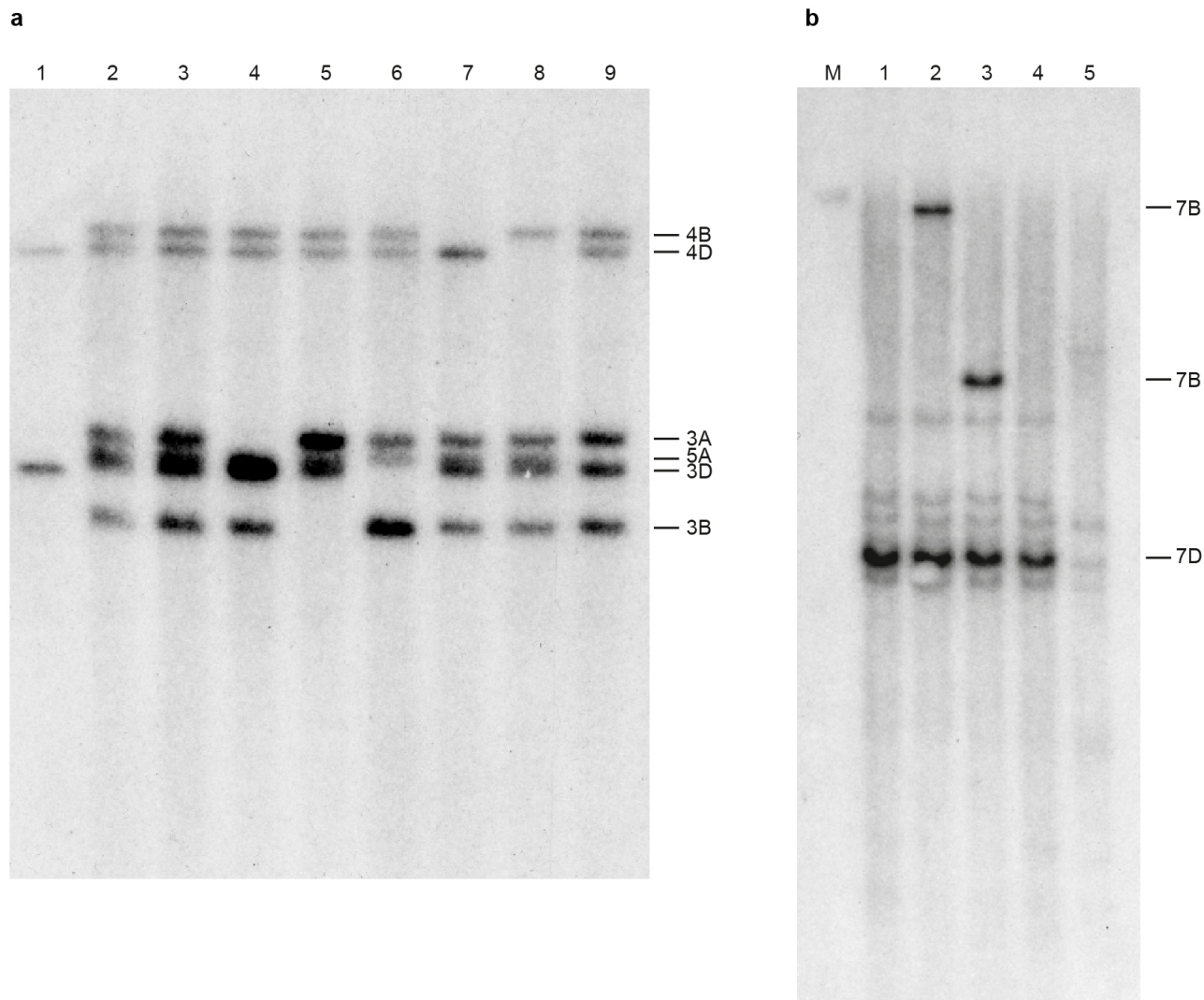
**Phylogenetic analysis of plant boron transporter genes sequences.** DNA sequences were either identified in this study or obtained from gene sequence databases<sup>23,35,36</sup> and the International Wheat Genome Sequencing Consortium (<http://www.wheatgenome.org>). Sequences were trimmed to cover the complete ORF. *Triticum urartu* and *Ae. tauschii* sequences are 99–100% identical to orthologous bread wheat genes and were not included in the phylogenetic analysis, with the exception of *AetBot-D5b*. Similarly, sequences of *Bot-B5c* and *Bot(Tp4A)-B5c* are nearly identical to that of *Bot-B5b* and not included. Nomenclature is in accordance with internationally accepted guidelines for wheat gene nomenclature and symbolization<sup>37</sup>. The rice locus LOC\_Os01g08020 contains two boron transporters of 96% sequence identity but is currently annotated as a single gene comprising a combination of the two transporters. Using transcript support from DQ421408 and AK072421 we generated two putative ORF sequences designated LOC\_Os01g08020\_gene A and LOC\_Os01g08020\_gene B. Phylogenetic analysis was performed using MEGA5 (ref. 38). Sequences were aligned using MUSCLE, and the maximum likelihood based on the Kimura 2 parameter model was calculated. To model evolutionary rates among sites, a discrete gamma distribution (two categories) was used ( $\Gamma = 1.1336$ ). The tree with

the highest log likelihood (−18,743.4333) is shown in Extended Data Fig. 4a. Bootstrap values were generated from 1,000 replicates.

**Transient expression in onion epidermal cells.** *Bot-B5b* coding sequence was cloned into the vector pMDC83 to generate a 35S:*Bot-B5b:GFP* (C terminus) construct. The 35S:*Bot-B5b:GFP* construct was used to transform onion (*Allium cepa*) epidermal cells by particle bombardment, and cells were visualized by confocal image analysis, as described previously<sup>39</sup>. Plasmolysis of onion epidermal cells was performed by immersion in 1 M sucrose for 1 min before image analysis.

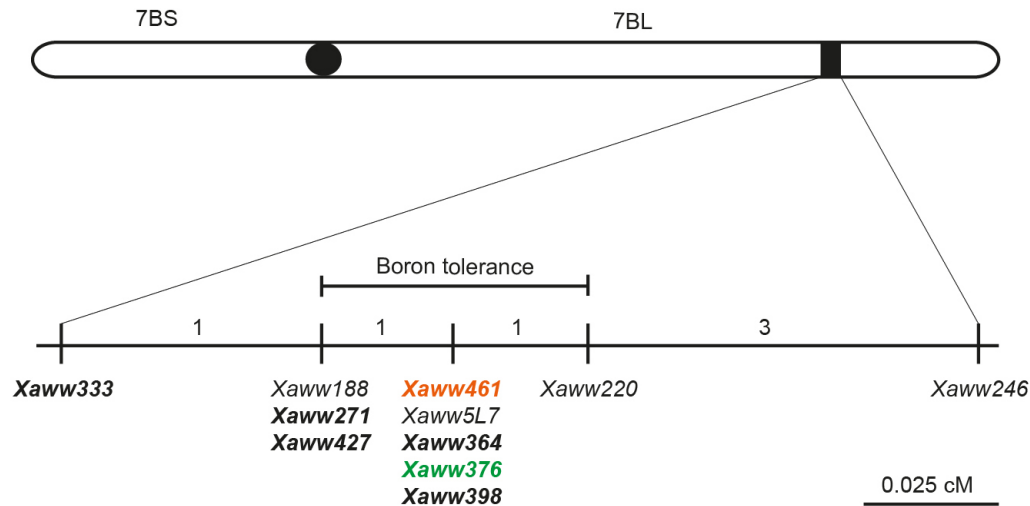
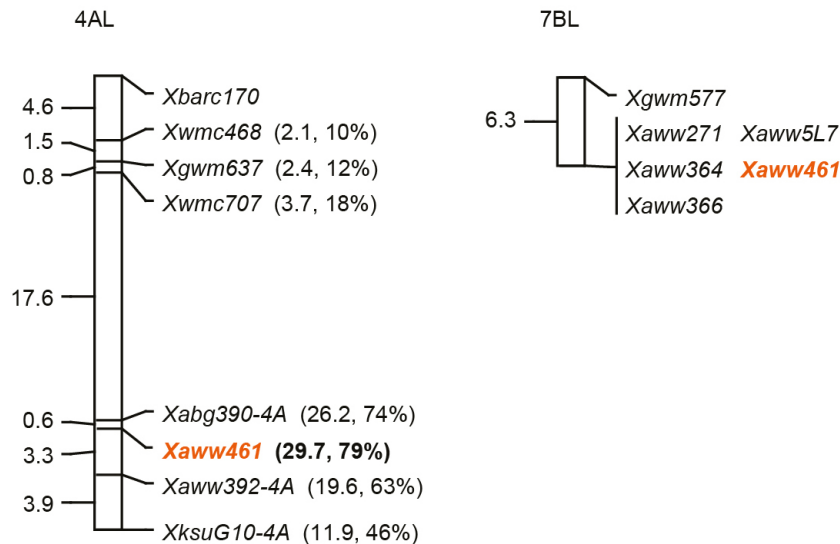
**Statistics.** Unless otherwise described, data were analysed using one-way analysis of variance (ANOVA) and Tukey's multiple comparisons tests ( $\alpha = 0.05$ ) for differences between treatments, using GraphPad Prism 6 software. qRT-PCR data of *Bot-B5* transcript levels (Fig. 2d) were log<sub>10</sub>-transformed before analysis. For analysis of root growth of genotypes across different boron concentrations shown in Fig. 1a, data for each genotype between 4 and 10 mM boron were compared by linear regression and Tukey's post-hoc testing for significant differences between slopes. For analysis of yeast growth (Fig. 2c), optical density data were plotted and fitted to Boltzmann sigmoidal functions by using nonlinear modelling with GraphPad Prism, to calculate times to half-saturation. Times to half-saturation ( $n = 5$ ) were then compared by one-way ANOVA and Tukey's Honestly Significant Difference test.

27. Chantachume, Y., Smith, D., Hollamby, G. J., Paull, J. G. & Rathjen, A. J. Screening for boron tolerance in wheat (*Triticum aestivum* L.) by solution culture in filter paper. *Plant Soil* **177**, 249–254 (1995).
28. Voorrips, R. E. MapChart: software for the graphical presentation of linkage maps and QTLs. *J. Hered.* **93**, 77–78 (2002).
29. Pallotta, M. A., Graham, R. D., Langridge, P., Sparrow, D. H. B. & Barker, S. J. RFLP mapping of manganese efficiency in barley. *Theor. Appl. Genet.* **101**, 1100–1108 (2000).
30. Pallotta, M. A. et al. in *Proceedings of the 10th International Wheat Genetics Symposium, Paestum, Italy, 1–6 September 2003* (ed. Pogna, N. E.) 789–791 (Istituto Sperimentale per la Cerealicoltura, 2003).
31. Burton, R. A., Shirley, N. J., King, B. J., Harvey, A. J. & Fincher, G. B. The *CesA* gene family of barley. Quantitative analysis of transcripts reveals two groups of co-expressed genes. *Plant Physiol.* **134**, 224–236 (2004).
32. Schreiber, A. et al. Comparative transcriptomics in the Triticeae. *BMC Genomics* **10**, 285 (2009).
33. Lai, K. et al. WheatGenome.info: an integrated database and portal for wheat genome information. *Plant Cell Physiol.* **53**, e2 (2012).
34. Gietz, R. D. & Woods, R. A. Transformation of yeast by the LiAc/SS carrier DNA/PEG method. *Methods Enzymol.* **350**, 87–96 (2002).
35. Jia, J. et al. *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature* **496**, 91–95 (2013).
36. Ling, H.-Q. et al. Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature* **496**, 87–90 (2013).
37. McIntosh, R. A. et al. *Catalogue of Gene Symbols for Wheat* <<http://wheat.pw.usda.gov/GG2/Triticum/wgc/2008/>> (2008).
38. Tamura, K. et al. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739 (2011).
39. Preuss, C. P., Huang, C. Y., Gilliam, M. & Tyerman, S. D. Channel-like characteristics of the low-affinity barley phosphate transporter PHT1;6 when expressed in *Xenopus* oocytes. *Plant Physiol.* **152**, 1431–1441 (2010).
40. Salse, J. et al. Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution. *Plant Cell* **20**, 11–24 (2008).



**Extended Data Figure 1 | Southern analysis of *Bot-D2a* and *Bot-B5* genes.**  
**a**, Genomic DNA from Chinese Spring nullisomic-tetrasomic (N-T) chromosome substitution lines digested with *Dra*I and hybridized with the probe AWW469, a 261-bp cDNA fragment of *Bot-D2* amplified from Cranbrook, which does not cross-hybridize to *Bot-B5* or *Bot-D5* at high stringency. Lane 1, *Ae. tauschii*; lane 2, Chinese Spring; lane 3, CS N7B-T7D; lane 4, CS N3A-T3D; lane 5, CS N3B-T3A; lane 6, CS N3D-T3B; lane 7, CS N4B-T4D; lane 8, CS N4D-T4A; lane 9, CS N5A-T5B. The chromosomal

location of each detected fragment is indicated. **b**, Genomic DNA was digested with *Hind*III and hybridized with the probe AWW471, a 357-bp genomic DNA fragment of *Bot-D5b* that hybridizes to *Bot-B5* at high stringency. Lane M, *Hind*III-digested Lambda DNA as size marker; lane 1, Cranbrook; lane 2, Halberd; lane 3, Chinese Spring; lane 4, CS nulli-tetra line N7B-T7A; lane 5, Langdon. The chromosomal location of strongly detected fragments is indicated. Minor bands are the result of low-level hybridization to paralogous sequences on 3A, 3B, 3D, 4B, 4D and 5A.

**a****b**

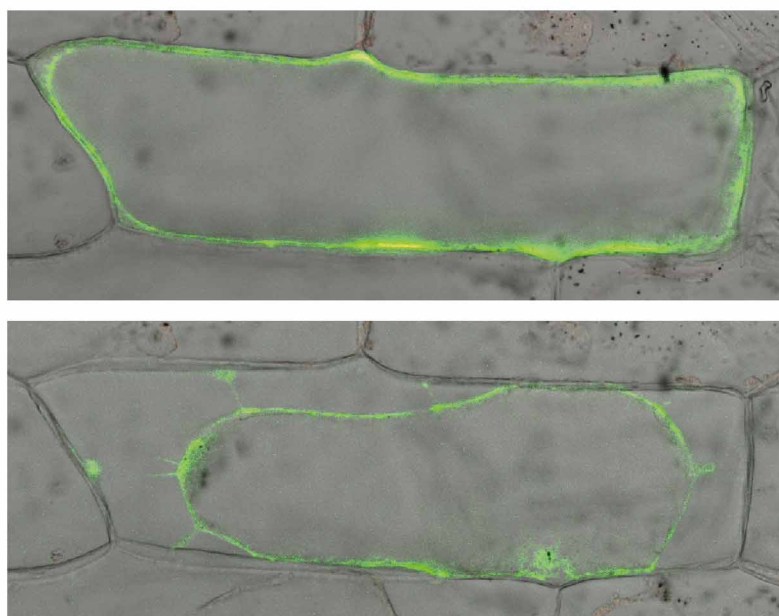
**Extended Data Figure 2 | Mapping of *Bot-B5* to 7BL (*Bo1*) and 4AL (*Bo4*).** **a**, Fine mapping of the *Bo1* locus on chromosome 7BL in a Cranbrook  $\times$  Halberd  $F_2$  population. Markers are listed below the line, and numbers above indicate recombinants identified for each marker interval. Previously unpublished markers are indicated in bold (see Supplementary Table 1). Markers in black font are derived from genes that are syntenous in rice, *Brachypodium* and wheat. The marker in green font is derived from a gene that is absent in rice but syntenous in *Brachypodium* and wheat. Marker AWW461 (orange font) is a fragment of *Bot-D5b* and is absent in both rice and *Brachypodium*. In both rice and *Brachypodium* the genes that are syntenous with AWW220 and AWW246 are immediately adjacent to each other, whereas

large insertion or inversion events have separated these genes in wheat.

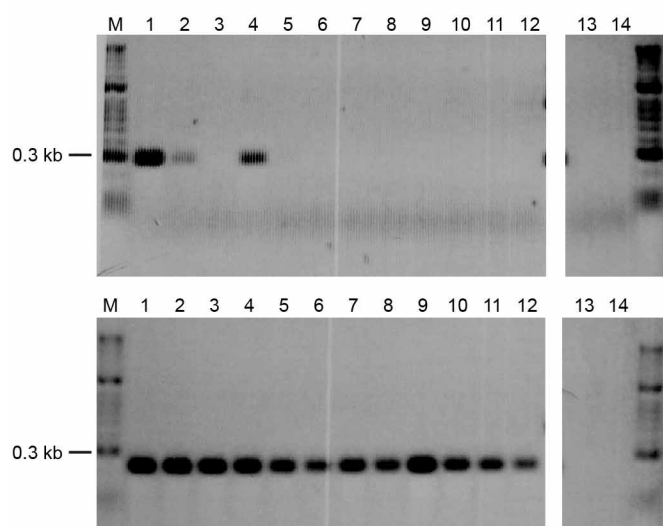
**b**, Partial genetic maps of chromosomes 4AL and 7BL in a G61450  $\times$  Kenya Farmer RIL population showing markers closely associated with *Bot-B5* alleles. Genetic distances in centimorgans are shown on the left side of the chromosome between markers. In brackets after the marker name we show LOD score and the percentage of total trait variation explained by the marker as derived using simple marker regression analysis for absolute longest root length in high-boron hydroponics. Data are shown for markers with a LOD score of  $>2.0$ . No significant marker-trait association was detected at the 7BL locus in this population.



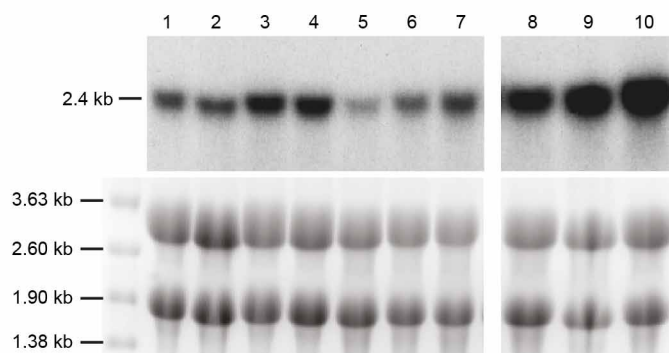
a



b



c



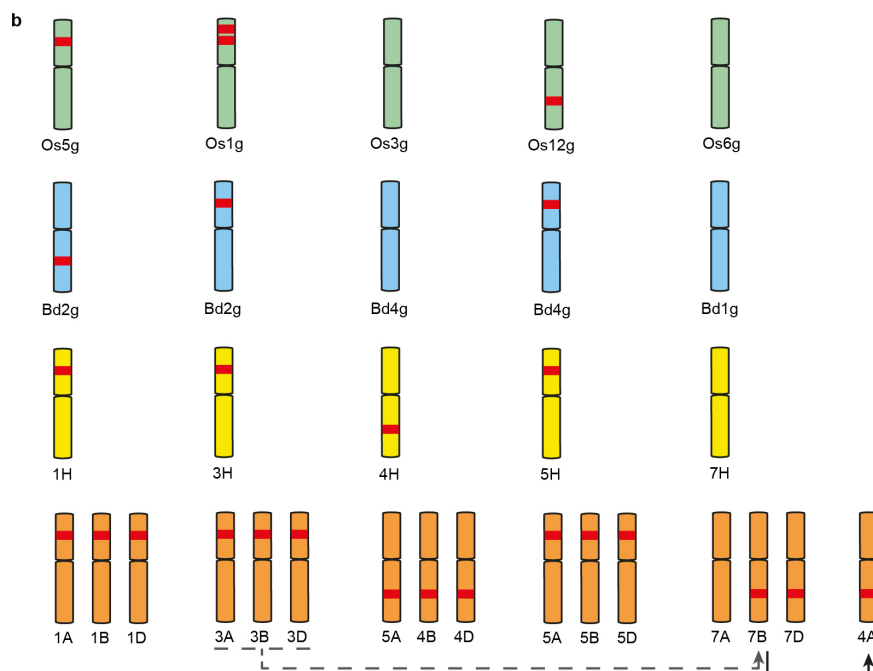
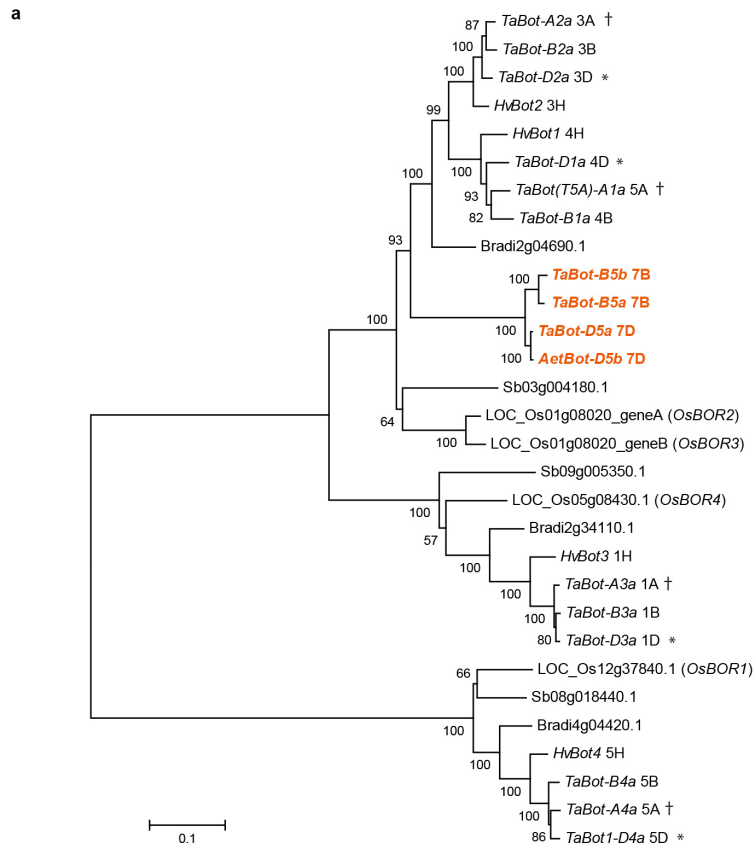
### Extended Data Figure 3 | *Bot-B5* is localized to the plasma membrane and is responsive to boron.

**a**, Confocal image of an onion epidermal cell with transient expression of 35S:*Bot-B5b::GFP* fusion protein before (upper panel) and after (lower panel) partial plasmolysis. Expression is confined to the plasma membrane. The lower panel shows Hechtian strands (thin membrane strands) connecting the cell wall to the plasmalemma after plasmolysis. The apparent signal in the sides of the cell in the upper panel is the result of a broad optical section imaged by the confocal microscope. This has resulted in an apparent broad distribution of green fluorescent protein (GFP) signal due to the curvature in cell wall towards the cell surface within the focal plane captured.

**b**, Semi-quantitative RT-PCR in a Chinese Spring developmental series<sup>32</sup> using the *Bot-B5*-specific marker qRT-PCR-*Bot-B5* (36 cycles, upper panel) and a wheat GAPDH marker (28 cycles, lower panel), indicating specific expression of *Bot-B5* in root tissues. Lane M, 2.5 µl of HyperLadder II from Bioline (Aust) Pty. Ltd; lane 1, root from 2-day-old germinating seeds; lane 2, embryo from 2-day-old germinating seeds; lane 3, coleoptiles from 2-day-old germinating seeds; lane 4, root from seedlings with shoots 10 cm long; lane 5, crown from seedlings with shoots 10 cm long; lane 6, leaf from seedlings with shoots 10 cm long; lane 7, immature inflorescences 2–3 cm long; lane 8, floral bracts 2 days before anthesis; lane 9, anthers 2 days before anthesis; lane 10, caryopsis

3–5 days after pollination; lane 11, embryo 22 days after pollination; lane 12, endosperm 22 days after pollination; lanes 13 and 14, water controls. A low level of product visible in the dissected embryo (lane 2) was probably due to contamination from root tissue.

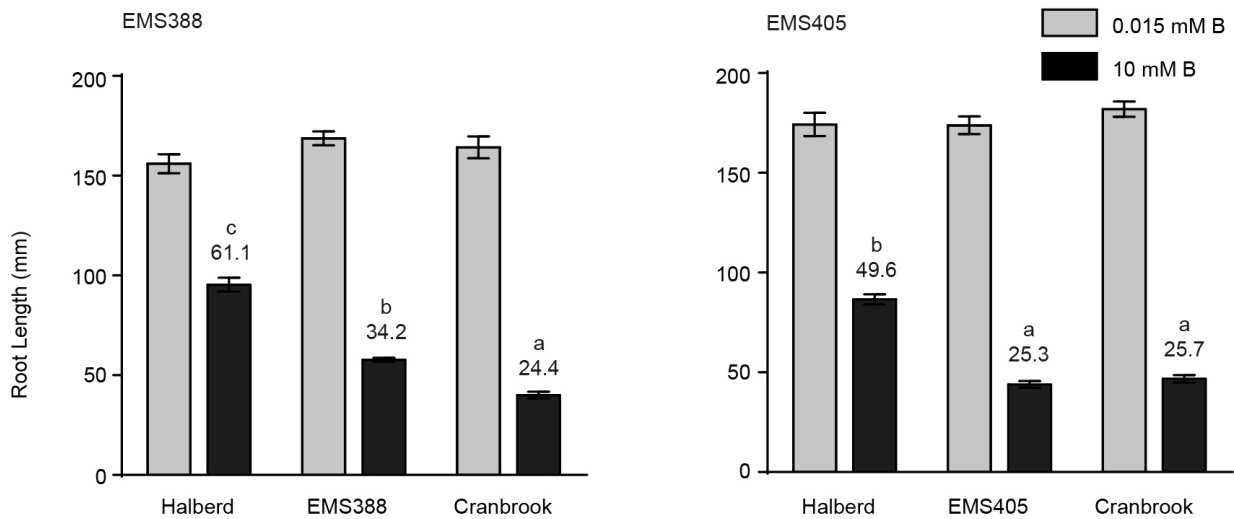
**c**, Northern analysis of *Bot-B5* mRNA levels in whole roots from 17-day-old Halberd and G61450 plants from hydroponics experiments 1 and 2 (see Methods). Seedlings were grown in nutrient solution for 17 days and treated either with no supplementary boron or with 2 mM supplementary boron for 22 h or 7 days before tissue collection. Lanes 1 and 2, Halberd plants from experiment 2 with no supplementary boron (two independent Halberd replicates are loaded); lanes 3 and 4, Halberd plants grown in experiment 2 and treated with 2 mM supplementary boron for 22 h (two independent Halberd replicates are loaded); lanes 5–7, Halberd plants grown in experiment 1 and treated with no supplementary boron, 2 mM supplementary boron for 22 h, and 2 mM supplementary boron for 7 days, respectively; lanes 8–10, G61450 plants grown in experiment 1 and treated with no supplementary boron, 2 mM supplementary boron for 22 h, and 2 mM supplementary boron for 7 days, respectively. Total RNA was revealed with ethidium bromide to indicate loading (lower panel) and analysed by northern hybridization (upper panel) using a 268-bp cDNA probe (AWW548) derived from Halberd, comprising 65 bp of coding sequence and 203 bp of 3' UTR.



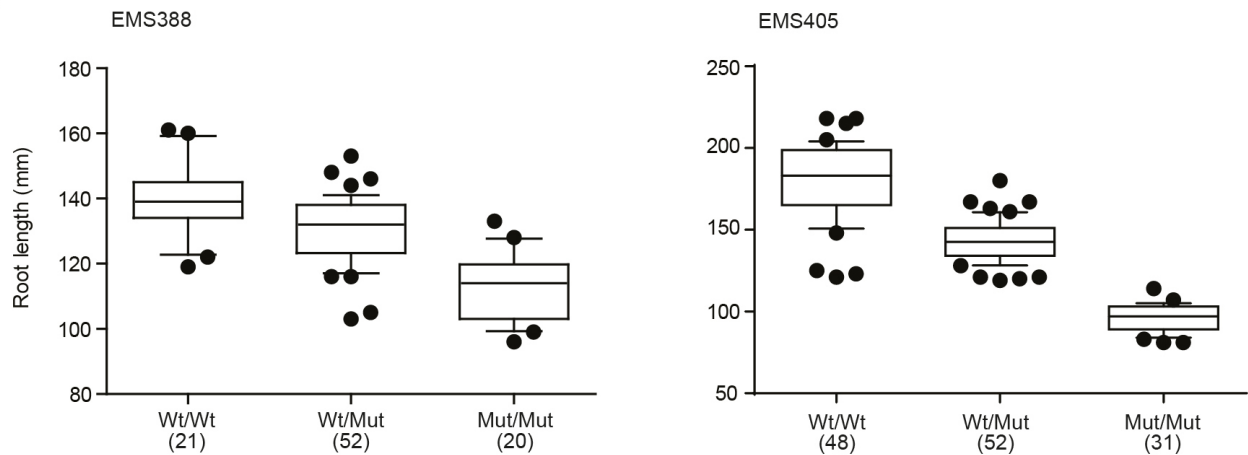
**Extended Data Figure 4 | Phylogeny and genome organization of monocot boron transporter genes.** **a**, Unrooted phylogenetic tree of boron transporter ORF sequences in selected monocot species (*Ta*, *T. aestivum*; *Aet*, *Ae. tauschii*; *Hv*, *Hordeum vulgare*; *Bradi*, *B. distachyon*; *Sb*, *S. bicolor*; *LOC\_Os*, *O. sativa*). Local bootstrap values (1,000 replicates) are shown as percentages adjacent to the branch line. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. *TaBot-B5* and *TaBot-D5* genes are indicated in orange font. Orthologous sequences in *Triticum uratu* are represented by a

dagger. As in bread wheat, no 7A gene is present in *T. uratu*. In addition to *AetBot-D5b*, orthologous sequences in *Ae. tauschii* are represented by an asterisk. **b**, Chromosomes of rice (green), *Brachypodium* (blue), barley (yellow) and wheat (orange), showing the approximate location of boron transporter sequences (red boxes) and grouped vertically by macro syntenic groups. The dispersed duplication of *Bot-B5* from 7BL to 4AL is shown by the solid arrow, and a putative ancient duplication between group 3 and group 7 chromosomes in wheat<sup>40</sup> is shown by the grey dashed arrow.

a

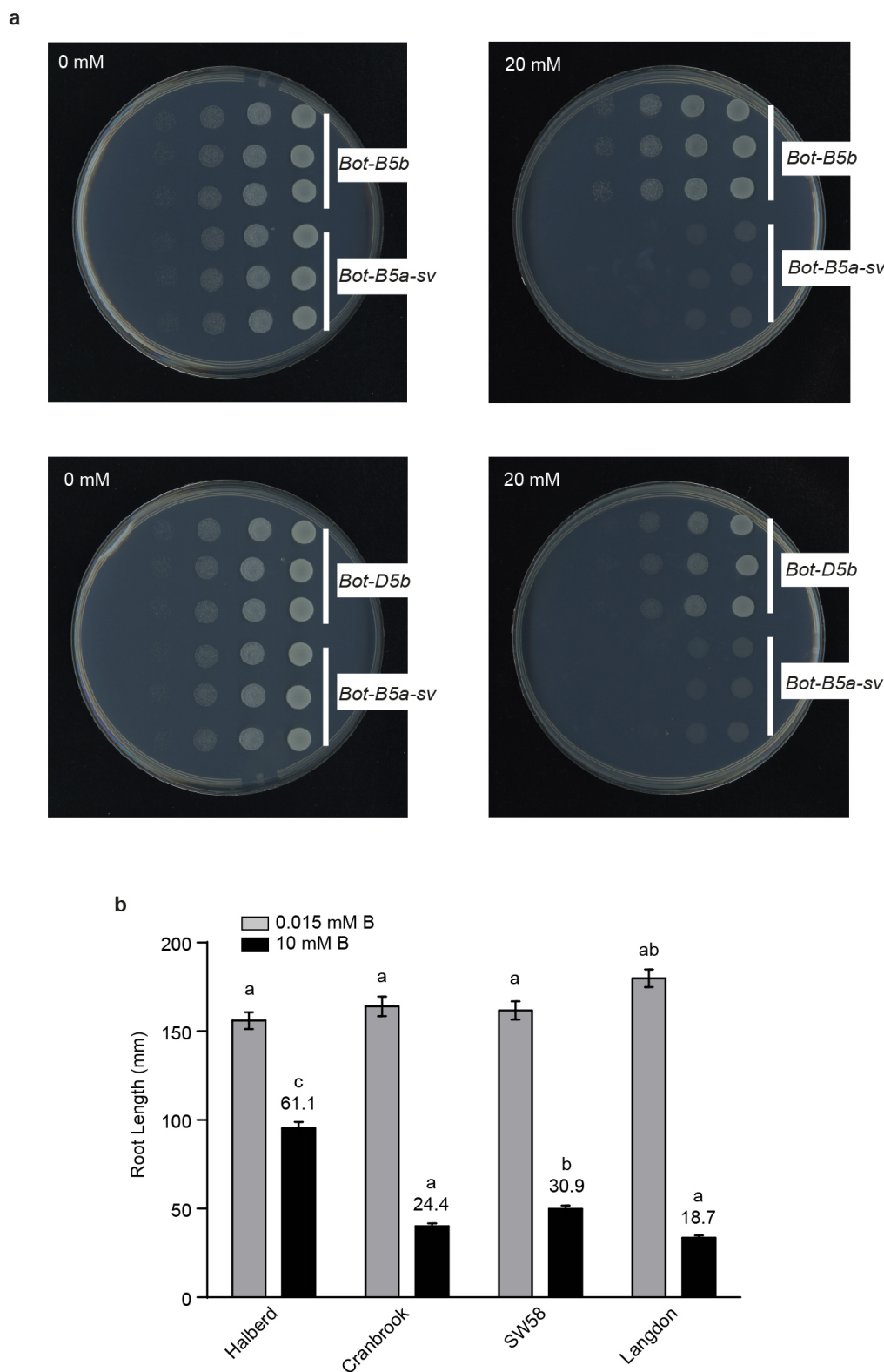


b



**Extended Data Figure 5 | Mutations in *Bot-B5b* reduce root growth at high boron.** **a**, Length of the longest root of wheat EMS mutant lines EMS388 and EMS405 in 0.015 mM boron and 10 mM boron in comparison with the standard cultivars Cranbrook and Halberd ( $n = 16$ ; means  $\pm$  s.e.m.). Numbers above the black columns are relative root lengths (root length in 10 mM boron expressed as a percentage of the root length in 0.015 mM boron). Letters denote significant ( $P < 0.01$ ) differences between genotypes at high boron. There were no genotypic differences at low boron. **b**, Root length of plants segregating for

mutant *Bot-B5b* alleles. Box plots show root lengths of seedlings segregating for *Bot-B5b-EMS388* ( $n = 93$ ) and *Bot-B5b-EMS405* ( $n = 131$ ). The boundaries of the boxes indicate 75th and 25th centiles, and lines within mark the median. Bars above and below the boxes indicate 90th and 10th centiles; outliers are shown as black circles. The longest root of a minimum of 20 individuals was measured for each group after hydroponic culture supplemented with 8 mM boron for 8 days (EMS388) or 13 days (EMS405). Numbers of plants measured for each allele class are indicated in parentheses.



**Extended Data Figure 6 | Analysis of protein function and boron tolerance of AeBot-D5b from synthetic wheat SW58.** **a**, Serial dilutions of yeast expressing *Bot-B5b* from Halberd (upper panels) or *Bot-D5b* from SW58 (lower panels), grown on solid medium containing no additional boron (left panels) and 20 mM supplementary boron (right panels). Each plate shows three independent yeast clones expressing either *Bot-B5b* or *Bot-D5b* at the top of the plate, and three independent clones expressing a truncated non-functional

sequence (*Bot-B5a-sv*) at the bottom of the plate. Aliquots of 10  $\mu$ l of tenfold serial dilutions of saturated cultures were spotted across the plates (right to left in each panel). **b**, Longest root length ( $n = 16$ ; means  $\pm$  s.e.m.) of 10-day-old seedlings of SW58, Halberd, Cranbrook and Langdon grown in 0.015 mM boron and 10 mM boron. Letters denote statistically distinct groups within each boron treatment (Tukey's Honestly Significant Difference test,  $\alpha = 0.05$ ).



Extended Data Table 1 | Wheat boron transporter nomenclature

Allele Name	Chromosome	Accession #	Alternate allele designation
<i>Bot(T5A)-A1a</i>	5AL		
<i>Bot-B1a</i>	4BL		
<i>Bot-D1a</i>	4DL		
<i>Bot-A2a</i>	3AS		
<i>Bot-B2a</i>	3BS		
<i>Bot-D2a</i>	3DS		<i>TaBor2</i> <sup>†</sup>
<i>Bot-A3a</i>	1AS		
<i>Bot-B3a</i>	1BS		
<i>Bot-D4a</i>	1DS		
<i>Bot-A4a</i>	5AS		<i>TaBOR1.2</i> <sup>†</sup>
<i>Bot-B4a</i>	5BS		<i>TaBOR1.3</i> <sup>†</sup>
<i>Bot-D4a</i>	5DS		<i>TaBOR1.1</i> <sup>†</sup>
<i>Bot-B5a</i>	7BL	KF148628	
<i>Bot-B5b</i>	7BL	KF148625	
<i>Bot(Tp4A)-B5c</i>	4AL	KF148626	
<i>Bot-B5c</i> <sup>‡</sup>	7BL	KF148627	
<i>Bot-B5d</i>	7BL	KF148629	
<i>Bot-B5e</i>	7BL	KF148630	
<i>Bot-B5f</i>	7BL	KF148631	
<i>Bot-B5g</i>	7BL	KF148633	
<i>Bot(Df)-B5h</i>	7BL	-	
<i>Bot-D5a</i>	7DL	KF148623	
<i>Bot-D5b</i>	7DL	KF148624	

Coding sequences and chromosomal locations of boron transporters on groups 1, 3, 4B/D and 5 were inferred from Chinese Spring genomic sequence obtained from the International Wheat Genome Sequencing Consortium chromosome survey sequences (<http://www.wheatgenome.org>). GenBank accession numbers of *BotB5/D5* allele sequences identified in this study are provided. *Df*, deficiency; *Tp*, transposition; *T*, translocation.

\* Sequence from line India 126 (synonym India) previously published as GenBank accession number EU220225 in ref. 14. The sequence is variant to the Chinese Spring sequence; however, we have not assigned a separate allele designation because the sequence differences have not been verified.

† Sequence from ref. 15.

‡ Accession KF148627 is derived from Lingzhi. *Bot-B5* promoter and coding sequence are identical for AUS10110a and AUS10344, differing from KF148627 by one single nucleotide polymorphism within the promoter at position 1,997 bp 5' of the ATG start site. Given the high level of sequence identity, the AUS10110a/AUS10344 7B allele is also named *Bot-B5c*.

# Parent-of-origin-specific allelic associations among 106 genomic loci for age at menarche

A list of authors and their affiliations appears at the end of the paper

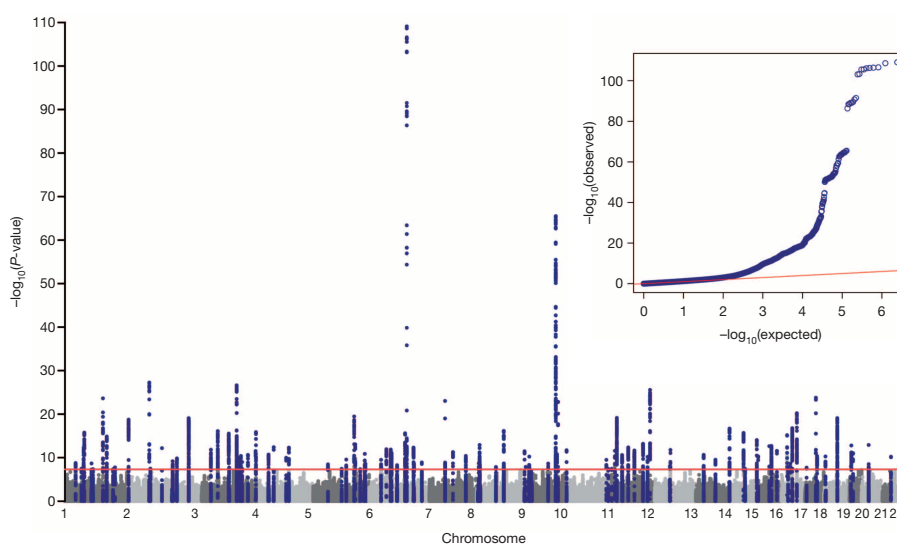
Age at menarche is a marker of timing of puberty in females. It varies widely between individuals, is a heritable trait and is associated with risks for obesity, type 2 diabetes, cardiovascular disease, breast cancer and all-cause mortality<sup>1</sup>. Studies of rare human disorders of puberty and animal models point to a complex hypothalamic-pituitary-hormonal regulation<sup>2,3</sup>, but the mechanisms that determine pubertal timing and underlie its links to disease risk remain unclear. Here, using genome-wide and custom-genotyping arrays in up to 182,416 women of European descent from 57 studies, we found robust evidence ( $P < 5 \times 10^{-8}$ ) for 123 signals at 106 genomic loci associated with age at menarche. Many loci were associated with other pubertal traits in both sexes, and there was substantial overlap with genes implicated in body mass index and various diseases, including rare disorders of puberty. Menarche signals were enriched in imprinted regions, with three loci (*DLK1-WDR25*, *MKRN3-MAGEL2* and *KCNK9*) demonstrating parent-of-origin-specific associations concordant with known parental expression patterns. Pathway analyses implicated nuclear hormone receptors, particularly retinoic acid and  $\gamma$ -aminobutyric acid-B2 receptor signalling, among novel mechanisms that regulate pubertal timing in humans. Our findings suggest a genetic architecture involving at least hundreds of common variants in the coordinated timing of the pubertal transition.

Genome-wide array data were available from up to 132,989 women of European descent from 57 studies. In a further 49,427 women, data were available on up to approximately 25,000 single nucleotide polymorphisms (SNPs), or their proxy markers, that showed sub-genome-wide

significant associations ( $P < 0.0022$ ) with age at menarche in our previous genome-wide association study (GWAS)<sup>4</sup> (Supplementary Table 1). Association statistics for 2,441,815 autosomal SNPs that passed quality control measures (including minor allele frequency  $> 1\%$ ) were combined across all studies by meta-analysis.

3,915 SNPs reached the genome-wide significance threshold ( $P < 5 \times 10^{-8}$ ) for association with age at menarche (Fig. 1). Using GCTA<sup>5</sup>, which approximates a conditional analysis adjusted for the effects of neighbouring SNPs (Extended Data Fig. 1 and Supplementary Table 2), we identified 123 independent signals for age at menarche at 106 genomic loci, including 11 loci containing multiple independent signals (Extended Data Tables 1–4; plots of all loci are available at <http://www.reprogen.org>). Of the 42 previously reported independent signals for age at menarche<sup>4</sup>, all but one (gene *SLC14A2*, SNP variation rs2243803,  $P = 2.3 \times 10^{-6}$ ) remained significant genome-wide in the expanded data set.

To estimate their overall contribution to the variation in age at menarche, we analysed an additional sample of 8,689 women. 104/123 signals showed directionally concordant associations or trends with menarche timing (binomial sign test  $P_{\text{Sign}} = 2.2 \times 10^{-15}$ ), of which 35 showed nominal significance ( $P_{\text{Sign}} < 0.05$ ) (Supplementary Table 3). In this independent sample, the top 123 SNPs together explained 2.71% ( $P < 1 \times 10^{-20}$ ) of the variance in age at menarche, compared to 1.31% ( $P = 2.3 \times 10^{-14}$ ) explained by the previously reported 42 SNPs. Consideration of further SNPs with lower levels of significance resulted in modest increases in the estimated variance explained with increasingly larger SNP sets, until we included all autosomal SNPs (15.8%, s.e. 3.6%,



**Figure 1 | Manhattan and quantile–quantile plot of the GWAS for age at menarche.** Manhattan (main panel) and quantile–quantile (QQ) (embedded) plots illustrating results of the genome-wide association study (GWAS) meta-analysis for age at menarche in up to 182,416 women of European descent. The Manhattan plot presents the association  $-\log_{10}(P\text{-values})$  for each genome-wide SNP (y axis) by chromosomal position (x axis). The red line

indicates the threshold for genome-wide statistical significance ( $P = 5 \times 10^{-8}$ ). Blue dots represent SNPs whose nearest gene is the same as that of the genome-wide significant signals. The QQ plot illustrates the deviation of association test statistics (blue dots) from the distribution expected under the null hypothesis (red line).

$P = 2.2 \times 10^{-6}$ ), indicating a highly polygenic architecture (Extended Data Fig. 2).

To test the relevance of menarche loci to the timing of related pubertal characteristics in both sexes, we examined their further associations with refined pubertal stage assessments in an overlapping subset of 10- to 12-year-old girls ( $n = 6,147$ ). A further independent sample of 3,769 boys had similar assessments at ages 12 to 15 years. 90/106 menarche loci showed consistent directions of association with Tanner stage in boys and girls combined ( $P_{\text{Sign}} = 1.1 \times 10^{-13}$ ), 86/106 in girls only ( $P_{\text{Sign}} = 6.2 \times 10^{-11}$ ) and 72/106 in boys only ( $P_{\text{Sign}} = 0.0001$ ), suggesting that the menarche loci are highly enriched for variants that regulate pubertal timing more generally (Supplementary Table 4).

Six independent signals were located in imprinted gene regions<sup>6</sup>, which is an enrichment when compared to all published genome-wide-significant signals for any trait and/or disease<sup>7</sup> (6/123, 4.8% vs 75/4332, 1.7%; Fisher's exact test  $P = 0.017$ ). Departure from Mendelian inheritance of pubertal timing has not been previously suspected, therefore we sought evidence for parent-of-origin-specific allelic associations in the deCODE Study, which included 35,377 women with parental origins of alleles determined by a combination of genealogy and long-range phasing<sup>6</sup>.

Two independent signals (no. 85a and 85b; rs10144321 and rs7141210) lie on chromosome 14q32 harbouring the reciprocally imprinted genes *DLK1* and *MEG3*, which exhibit paternal-specific or maternal-specific expression, respectively, and may underlie the growth retardation and precocious puberty phenotype of maternal uniparental disomy-14<sup>8</sup>. In deCODE, for both signals the paternally inherited alleles were associated with age at menarche (rs10144321,  $P_{\text{pat}} = 3.1 \times 10^{-5}$ ; rs7141210,  $P_{\text{pat}} = 2.1 \times 10^{-4}$ ), but the maternally inherited alleles were not ( $P_{\text{mat}} = 0.47$  and  $0.12$ , respectively), and there was significant heterogeneity between paternal and maternal effect estimates (rs10144321,  $P_{\text{het}} = 0.02$ ; rs7141210,  $P_{\text{het}} = 2.2 \times 10^{-4}$ ) (Fig. 2; Supplementary Table 5). Notably, rs7141210 is reportedly a *cis*-acting methylation-quantitative trait locus

(QTL) in adipose tissue<sup>9</sup> (Extended Data Table 5) and the menarche age-raising allele was also associated with lower transcript levels of *DLK1* (Supplementary Tables 6 and 7)<sup>10</sup>, which encodes a transmembrane protein involved in adipogenesis and neurogenesis. In deCODE data, the maternally inherited rs7141210 allele was correlated with blood transcript levels of the maternally expressed genes *MEG3* ( $P_{\text{mat}} < 5.6 \times 10^{-53}$ ), *MEG8* ( $P_{\text{mat}} = 4.9 \times 10^{-41}$ ) and *MEG9* ( $P_{\text{mat}} = 5.4 \times 10^{-5}$ ); however, lack of any correlation with the paternally inherited alleles ( $P_{\text{pat}} = 0.18$ ,  $P_{\text{pat}} = 0.87$  and  $P_{\text{pat}} = 0.37$ , respectively) suggests that these genes do not explain this paternal-specific menarche signal.

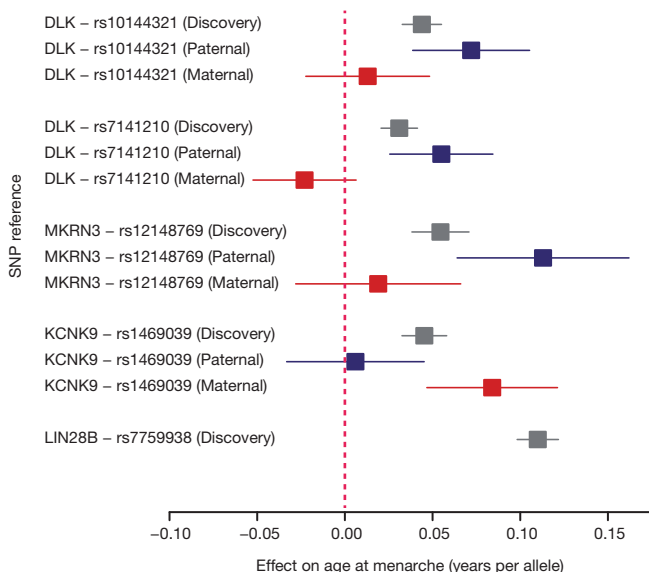
Signal no. 86 (rs12148769) lies in the imprinted critical region for Prader-Willi syndrome, which is caused by paternal-specific deletions of chromosome 15q11-13 and includes clinical features of hypogonadotropic hypogonadism and hypothalamic obesity<sup>11</sup>; conversely, a small proportion of cases have precocious puberty. For rs12148769, only the paternally inherited allele was associated with age at menarche ( $P_{\text{pat}} = 2.4 \times 10^{-6}$ ), but the maternally inherited allele was not ( $P_{\text{mat}} = 0.43$ ;  $P_{\text{het}} = 5.6 \times 10^{-3}$ ) (Fig. 2). Recently, truncating mutations of *MAGEL2* affecting the paternal alleles were reported in Prader-Willi syndrome; all four reported cases had hypogonadism or delayed puberty<sup>11</sup>, whereas paternally inherited deleterious mutations in *MKRN3* were found in patients with central precocious puberty<sup>3</sup>. It is as yet unclear which of these paternally expressed genes explains this menarche signal.

Signal no. 57 (rs1469039) is intronic in *KCNK9*, which shows maternal-specific expression in mouse and human brain<sup>12</sup>. Concordantly, only the maternally inherited allele was associated with age at menarche ( $P_{\text{mat}} = 5.6 \times 10^{-6}$ ), but the paternally inherited allele was not ( $P_{\text{pat}} = 0.76$ ;  $P_{\text{het}} = 3.7 \times 10^{-3}$ ) (Fig. 2). The menarche age-increasing allele was associated with lower transcript levels of *KCNK9* in deCODE's blood expression data when maternally inherited ( $P_{\text{mat}} = 0.003$ ), but not when paternally inherited ( $P_{\text{pat}} = 0.31$ ). *KCNK9* encodes TASK-3, which belongs to a family of two-pore domain potassium channels that regulate neuronal resting membrane potential and firing frequency.

The two remaining signals located within imprinted regions (rs2137289 and rs947552) did not demonstrate either paternal- or maternal-specific association. We then systematically tested all 117 remaining independent menarche signals for parent-of-origin-specific associations with menarche timing and found only four (3.4%) with at least nominal associations ( $P_{\text{het}} < 0.05$ ; Supplementary Table 5), which was proportionately fewer than signals at imprinted regions (4/6 (67.0%), Wilcoxon rank sum test  $P = 0.009$ ).

Three menarche signals were in genes encoding JmjC-domain-containing lysine-specific demethylases (enrichment  $P = 0.006$  for all genes in this family); signal no. 1 (rs2274465) is intronic in *KDM4A*, signal no. 37 (rs17171818) is intronic in *KDM3B*, and signal no. 59b (rs913588) is a missense variant in *KDM4C*. Notably, *KDM3B*, *KDM4A* and *KDM4C* all encode activating demethylases for lysine 9 on histone H3, which was recently identified as the chromatin methylation target that mediates the remarkable long-range regulatory effects of *IPW*, a paternally expressed long noncoding RNA in the imprinted Prader-Willi syndrome region on chromosome 15q11-13, on maternally expressed genes at the imprinted *DLK1-MEG3* locus on chromosome 14q32<sup>13</sup>. Examination of sub-genome-wide signals showed another potential locus intronic in *KDM4B* (rs11085110,  $P = 2.3 \times 10^{-6}$ ). Pubertal onset in female mice is reportedly triggered by DNA methylation of the Polycomb group silencing complex of genes (including *CBX7* near signal no. 105), leading to enrichment of activating lysine modifications on histone H3<sup>14</sup>. Specific histone demethylases could potentially regulate cross-links between imprinted regions to influence pubertal timing.

Menarche signals also tended to be enriched in or near genes that underlie rare Mendelian disorders of puberty (enrichment  $P = 0.05$ )<sup>2,3</sup>. As well as rs12148769 near *MKRN3*, signals were found near *LEPR-LEPROT* (signal no. 2; rs10789181), which encodes the leptin receptor, and immediately upstream of *TACR3* (signal no. 32; rs3733631), which encodes the receptor for neurokinin B. A further variant approximately 10 kilobases (kb) from *GNRH1* approached genome-wide significance (rs1506869,



**Figure 2 | Forest plot of parent-of-origin-specific allelic associations at three imprinted menarche loci.** The forest plot illustrates the associations of variants in four independent genomic signals for age at menarche that are located in three imprinted gene regions. For each variant, squares (and error bars) indicate the estimated per-allele effect sizes on age at menarche in years (and 95% confidence intervals) from the standard additive models in the combined ReproGen meta-analysis (grey), and separately for the paternally inherited (blue) or maternally inherited allele (red) in up to 35,377 women from the deCODE study. The association for the menarche locus with the largest effect size at *LIN28B* is also shown for reference, illustrating the similar magnitude of effect size at the *MKRN3* locus when parent-of-origin is taken into account.

$P = 1.8 \times 10^{-6}$ ) and was also associated with *GNRH1* expression in adipose tissue ( $P = 3.7 \times 10^{-5}$ ). Signals no. 34 (rs17086188) and 103 (rs852069) lie near *PCSK1* and *PCSK2*, respectively, indicating a common function of the type 1 and 2 prohormone convertases in pubertal regulation. Signals in or near several further genes with relevance to pituitary development/function included: signal no. 20 (rs7642134) near *POU1F1*, signal no. 39 (rs9647570) within *TENM2*, and signal no. 42 (rs2479724) near *FRS3*. Furthermore, signals no. 71 (rs7103411) and no. 92 (rs1129700) are *cis*-expression QTLs (eQTLs) for *LGR4* and *TBX6*, respectively, both of which encode enhancers for the pituitary development factor *SOX2*. Signals no. 52 (rs6964833) intronic in *GTF2I* and no. 104 (rs2836950) intronic in *BRWD1* were found in critical regions for complex conditions that include abnormal reproductive phenotypes, Williams–Beuren syndrome (early puberty)<sup>15</sup> and Down syndrome (hypogonadism in boys), respectively<sup>16</sup>.

Including signals described above, we identified 29 menarche signals in or near genes with possible roles in hormonal functions (Fig. 3, Supplementary Table 8), many more than the three signals we described previously (*INHBA*, *PCSK2* and *RXRG*)<sup>4</sup>. Two signals were found in or near genes related to steroidogenesis. Signal 35 (rs251130) was a *cis*-eQTL for *STARD4*, which encodes a StAR-related lipid transfer protein involved in the regulation of intra-cellular cholesterol trafficking. Signal no. 9 (rs6427782) is near *NR5A2*, which encodes a nuclear receptor with key roles in steroidogenesis and oestrogen-dependent cell proliferation.

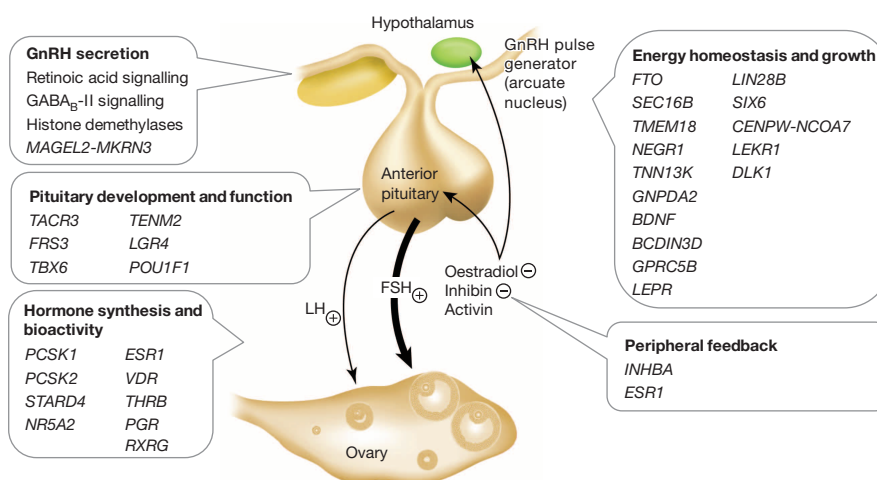
We observed that SNPs in or near a custom list of genes that encode nuclear hormone receptors, co-activators or co-repressors were enriched for associations with menarche timing (enrichment  $P = 6 \times 10^{-5}$ ). Individually, nine genome-wide significant signals mapped to within 500 kb of these genes, including those encoding the nuclear receptors for oestrogen, progesterone, thyroid hormone and 1,25-dihydroxyvitamin D3. Several nuclear hormone receptors are involved in retinoic acid signalling. SNPs in or near *RXRG* and *RORA* reached genome-wide significance, and three other genes contained sub-genome-wide signals (*RXRA* (rs2520094,  $P = 4 \times 10^{-7}$ ), *RORB* (rs4237264,  $P = 9.4 \times 10^{-6}$ ), *RXRB* (rs241438,  $P = 7.1 \times 10^{-5}$ )). Two other genome-wide significant signals mapped to genes with roles in retinoic acid function (no. 67 *CTBP2* and no. 101 *RDH8*). The active metabolites of vitamin A, all-*trans*-retinoic acid and 9-*cis*-retinoic acid, have differential effects on gonadotropin-releasing hormone (GnRH) expression and secretion<sup>17</sup>. Other possible mechanisms linking retinoic acid signalling to pubertal timing include inhibition of embryonic GnRH neuron migration, and enhancement of steroidogenesis and gonadotropin secretion<sup>18</sup>. The relevance of our findings to observations of low circulating vitamin A levels and use of dietary vitamin A in delayed puberty<sup>19</sup> are yet unclear.

To identify other mechanisms that regulate pubertal timing, we tested all SNPs genome-wide for collective enrichment across any biological pathway defined in publicly available databases. The top ranked pathway reaching study-wide significance (false discovery rate = 0.009) was gamma-aminobutyric acid (GABA<sub>B</sub>) receptor II signalling (Extended Data Table 6); each of the nine genes in this pathway contained a SNP with sub-genome-wide significant association with menarche (Extended Data Table 7). Notably, GABA<sub>B</sub> receptor activation inhibits hypothalamic GnRH secretion in animal models<sup>20</sup>.

Regarding the relevance of our findings to other traits, we confirmed<sup>4</sup> and extended the overlap between genome-wide significant loci for menarche and adult body mass index (BMI)<sup>21</sup>. At all nine loci (in or near *FTO*, *SEC16B*, *TMEM18*, *NEGR1*, *TNNI3K*, *GNPDA2*, *BDNF*, *BCDIN3D* and *GPRC5B*) the menarche age-raising allele was also associated with lower adult BMI (Supplementary Table 9). Three menarche signals overlapped known loci for adult height<sup>22</sup>. The menarche age-raising alleles at signals no. 47c (rs7759938, *LIN28B*) and no. 83 (rs1254337, *SIX6*) were also associated with taller adult height, which is directionally concordant with epidemiological observations. Conversely, the menarche age-raising allele at signal no. 48 (rs4895808, *CENPW-NCOA7*) was associated with shorter adult height (Supplementary Table 9).

Further menarche signals overlapped reported GWAS loci for other traits, but in each case at only a single locus, therefore possibly reflecting small-scale pleiotropy rather than a broader shared genetic aetiology. Signal no. 26 (rs900400) was a *cis*-eQTL for *LEKR1*, and is the same lead SNP associated with birth weight<sup>23</sup>. The menarche age-raising allele was also associated with higher birth weight, directionally concordant with epidemiological observations<sup>24</sup>. Signal no. 48 (rs4895808, a *cis*-eQTL for *CENPW*) is in linkage disequilibrium (LD) ( $r^2 = 0.90$ ) with the lead SNP for the autoimmune disorder type 1 diabetes, rs9388489<sup>25</sup>, which also showed robust association with menarche timing ( $P = 6.49 \times 10^{-12}$ ). Signal no. 41 (rs16896742) is near *HLA-A*, which encodes the class I, A major histocompatibility complex, and is a known locus for various immunity or inflammation-related traits<sup>7</sup>. Signal no. 50 (rs6933660) is near *ESR1*, which encodes the oestrogen receptor, a known locus for breast cancer<sup>26</sup> and bone mineral density<sup>27</sup>. Notably, the menarche age-raising allele at rs6933660 was associated with higher femoral neck bone mineral density ( $P = 6 \times 10^{-5}$ )<sup>27</sup>, which is directionally discordant with the epidemiological association<sup>28</sup>. Signal no. 70 (rs11022756) is intronic in *ARNTL*, a known locus for circulating plasminogen activator inhibitor type 1 (PAI-1) levels<sup>29</sup>; the reported lead SNP (rs6486122) for PAI-1<sup>29</sup> also showed robust association with menarche timing ( $P = 9.3 \times 10^{-10}$ ).

Our findings indicate both BMI-related and BMI-independent mechanisms that could underlie the epidemiological associations between



**Figure 3 | Schematic diagram indicating possible roles in the hypothalamic-pituitary-ovarian axis of several of the implicated genes and biological mechanisms for menarche timing.**



early menarche and higher risks of adult disease<sup>1</sup>. These include actions of *LIN28B* on insulin sensitivity through the mTOR pathway, GABA<sub>B</sub> receptor signalling on inhibition of oxidative stress-related  $\beta$ -cell apoptosis, and *SIRT3* (mitochondrial sirtuin 3), which could link early life nutrition to metabolism and ageing. Finally, only few parent-of-origin-specific allelic associations at imprinted loci have been described for complex traits<sup>6</sup>. Our findings implicate differential pubertal timing, a trait with putative selection advantages<sup>30</sup>, as a potential additional target for the evolution of genomic imprinting.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 23 December 2013; accepted 30 May 2014.

Published online 23 July 2014.

- Prentice, P. & Viner, R. M. Pubertal timing and adult obesity and cardiometabolic risk in women and men: a systematic review and meta-analysis. *Int. J. Obes.* **37**, 1036–1043 (2013).
- Silveira, L. F. G. & Latronico, A. C. Approach to the patient with hypogonadotropic hypogonadism. *J. Clin. Endocrinol. Metab.* **98**, 1781–1788 (2013).
- Abreu, A. P. et al. Central precocious puberty caused by mutations in the imprinted gene *MKRN3*. *N. Engl. J. Med.* **368**, 2467–2475 (2013).
- Elks, C. E. et al. Thirty new loci for age at menarche identified by a meta-analysis of genome-wide association studies. *Nature Genet.* **42**, 1077–1085 (2010).
- Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nature Genet.* **44**, 369–375 (2012).
- Kong, A. et al. Parental origin of sequence variants associated with complex diseases. *Nature* **462**, 868–874 (2009).
- Hindorf, L. A. et al. A catalog of published genome-wide association studies. Available at <http://www.genome.gov/gwastudies>. (Accessed, 1 November 2013).
- Temple, I. K., Shrubbs, V., Lever, M., Bullman, H. & Mackay, D. J. G. Isolated imprinting mutation of the *DLK1/GTL2* locus associated with a clinical presentation of maternal uniparental disomy of chromosome 14. *J. Med. Genet.* **44**, 637–640 (2007).
- Grundberg, E. et al. Global analysis of DNA methylation variation in adipose tissue from twins reveals links to disease-associated variants in distal regulatory elements. *Am. J. Hum. Genet.* **93**, 876–890 (2013), corrected **93**, 1158 (2013).
- Westra, H.-J. et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nature Genet.* **45**, 1238–1243 (2013).
- Schaefer, C. P. et al. Truncating mutations of *MAGEL2* cause Prader-Willi phenotypes and autism. *Nature Genet.* **45**, 1405–1408 (2013).
- Ruf, N. et al. Sequence-based bioinformatic prediction and QUASEP identify genomic imprinting of the *KCNK9* potassium channel gene in mouse and human. *Hum. Mol. Genet.* **16**, 2591–2599 (2007).
- Stelzer, Y., Sagi, I., Yanuka, O., Eiges, R. & Benvenisty, N. The noncoding RNA *IPW* regulates the imprinted *DLK1-DIO3* locus in an induced pluripotent stem cell model of Prader-Willi syndrome. *Nature Genet.* **46**, 551–557 (2014).
- Lomniczi, A. et al. Epigenetic control of female puberty. *Nature Neurosci.* **16**, 281–289 (2013).
- Partsch, C.-J. et al. Central precocious puberty in girls with Williams syndrome. *J. Pediatr.* **141**, 441–444 (2002).
- Grinspon, R. P. et al. Early onset of primary hypogonadism revealed by serum anti-Müllerian hormone determination during infancy and childhood in trisomy 21. *Int. J. Androl.* **34**, e487–e498 (2011).
- Cho, S. et al. 9-cis-Retinoic acid represses transcription of the gonadotropin-releasing hormone (GnRH) gene via proximal promoter region that is distinct from all-trans-retinoic acid response element. *Brain Res. Mol. Brain Res.* **87**, 214–222 (2001).
- Nagl, F. et al. Retinoic acid-induced nNOS expression depends on a novel PI3K/Akt/DAX1 pathway in human TGW-nu-1 neuroblastoma cells. *Am. J. Physiol. Cell Physiol.* **297**, C1146–C1156 (2009).
- Zadik, Z., Sinai, T., Zung, A. & Reifen, R. Vitamin A and iron supplementation is as efficient as hormonal therapy in constitutionally delayed children. *Clin. Endocrinol.* **60**, 682–687 (2004).
- Constantin, S. et al. GnRH neuron firing and response to GABA *in vitro* depend on acute brain slice thickness and orientation. *Endocrinology* **153**, 3758–3769 (2012).
- Speliotes, E. K. et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature Genet.* **42**, 937–948 (2010).
- Lango Allen, H. et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832–838 (2010).
- Horikoshi, M. et al. New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. *Nature Genet.* **45**, 76–82 (2013).
- D'Aloisio, A. A., DeRoo, L. A., Baird, D. D., Weinberg, C. R. & Sandler, D. P. Prenatal and infant exposures and age at menarche. *Epidemiology* **24**, 277–284 (2013).
- Barrett, J. C. et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nature Genet.* **41**, 703–707 (2009).
- Zheng, W. et al. Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nature Genet.* **41**, 324–328 (2009).
- Estrada, K. et al. Genome-wide meta-analysis identifies 56 bone mineral density loci and reveals 14 loci associated with risk of fracture. *Nature Genet.* **44**, 491–501 (2012).
- Parker, S. E. et al. Menarche, menopause, years of menstruation, and the incidence of osteoporosis: the influence of prenatal exposure to diethylstilbestrol. *J. Clin. Endocrinol. Metab.* **99**, 594–601 (2014).
- Huang, J. et al. Genome-wide association study for circulating levels of PAI-1 provides novel insights into its regulation. *Blood* **120**, 4873–4881 (2012).
- Migliano, A. B., Vinicius, L. & Lahr, M. M. Life history trade-offs explain the evolution of human pygmies. *Proc. Natl Acad. Sci. USA* **104**, 20216–20219 (2007).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** A full list of acknowledgements can be found in the Supplementary Information.

**Author Contributions** Overall project management: J.R.B.P., F.D., C.E.E., P.S., D.J.T., D.F.E., K.S., J.M.M. and K.K.O. Core analyses: J.R.B.P., F.D., C.E.E., P.S., T.F., D.J.T., D.I.C. and T.E. Individual study analysts: A.A.R., A.D., A.G., A.J., A.T., A.V.S., B.Z.A., B.F., C.E.E., D.F.G., D.I.C., D.J.T., D.L.C., D.L.K., E.A., E.K.W., E.M., E.M.B., E.T., F.D., G.M., G.McMahon, I.M.N., J.A.V., J.D., J.H., J.R.B.P., J.T., J.Z., K.L.L., K.M., L.L.P., L.M.R., L.M.Y., L.S., M.M., N.F., N.Ts., P.K., P.S., R.M., S.K., S.S., S.S.U., T.C., T.E., T.F., T.Fo., T.H.P., W.Q.A. and Z.K. Individual study data management and generation: A.A.R., A.C.H., A.D., A.D.C., A.G.U., A.J.O., A.M.S., A.Mu., A.P., A.Po., B.A.O., C.A.H., D.C., D.I.C., D.J.H., D.K., D.Lw., D.P.K., D.P.S., D.S., E.A.N., E.P., E.W., F.A., F.B.H., F.G., F.R., G.D., G.E., G.G.W., H.S., H.W., I.D., J.C., J.H., J.P.R., L.F., L.Fr., L.M., L.M.R., M.E.G., M.J.S., M.J.W., M.K.B., M.Melbye, M.P., M.W., N.A., N.J.T., N.L.P., P.K.M., Q.W., R.H., S.B., S.C., S.G., S.L., S.R., S.S.U., T.E., U.S., U.T., V.S. and W.L.M. Individual study principal investigators: A.C., A.G.U., A.H., A.J.O., A.K.D., A.L., A.M., A.M.D., A.Mannermaa, A.Mu., A.R., B.B., B.Z.A., B.H.R.W., C.B., C.E.P., C.G., C.H., C.van Duijn, D.I.B., D.F., D.F.E., D.J.H., D.L., D.Lw., D.S.P., D.P.S., D.Schlessinger, E.A.S., E.B., E.E.J.d.G., E.L., E.W., E.W.D., F.B.H., F.J.C., G.C., G.D., G.G.G., G.Wa., G.Wi., G.W.M., H.A., H.A.B., H.B., H.Be., H.F., H.N., H.S., H.V., I.D., I.L.A., J.A.K., J.B., J.C.C., J.G.E., J.E.B., J.L.H., J.M.C., J.M.M., J.P., K.C., K.K., K.K.O., K.P., K.S., L.C., L.F., L.J.B., M.C.S., M.G., M.I.M., M.J., M.J.E., M.J.H., M.J.S., M.K.S., M.W.B., M.Z., N.G.M., N.J.W., P.A.F., P.D., P.D.P.P., P.F.M., P.G., P.H., P.K., P.M.R., P.N., P.P., P.P.G., P.R., P.V., R.J.F.L., R.L.M., R.W., S.B., S.Bergmann, S.C., S.E.B., T.B.H., T.D.S., T.I.A.S., U.H., V.G., V.K. and V.S.

**Author Information** Plots of all 106 menarche loci and genome-wide summary level statistics are available at the ReproGen Consortium website: <http://www.reprogen.org>. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.R.B.P. (john.perry@mrc-epid.cam.ac.uk) and J.M. (marabito@bu.edu).

John R. B. Perry<sup>1,2,3,4\*</sup>, Felix Day<sup>1\*</sup>, Cathy E. Elks<sup>1\*</sup>, Patrick Sulem<sup>5\*</sup>, Deborah J. Thompson<sup>6</sup>, Teresa Ferreira<sup>3</sup>, Chunyan He<sup>7,8</sup>, Daniel I. Chasman<sup>9,10</sup>, Tõnu Esko<sup>11,12,13,14</sup>, Gudmar Thorleifsson<sup>5</sup>, Eva Albrecht<sup>15</sup>, Wei Q. Ang<sup>16</sup>, Tanguy Corre<sup>17,18</sup>, Diana L. Cousminer<sup>19</sup>, Bjarke Feenstra<sup>20</sup>, Nora Franceschini<sup>21</sup>, Andrea Ganna<sup>22</sup>, Andrew D. Johnson<sup>23</sup>, Sanela Kjellqvist<sup>24</sup>, Kathryn L. Lunetta<sup>23,25</sup>, George McMahon<sup>26,27</sup>, Ilja M. Nolte<sup>28</sup>, Lavinia Paternoster<sup>26</sup>, Eleonora Porcu<sup>29,30</sup>, Albert V. Smith<sup>31,32</sup>, Lisette Stolk<sup>33,34</sup>, Alexander Teumer<sup>35</sup>, Natalia Tšernikova<sup>11,36</sup>, Emmi Tikkanen<sup>19,37</sup>, Sheila Ulivi<sup>38</sup>, Erin K. Wagner<sup>7,8</sup>, Najaf Amin<sup>39</sup>, Laura J. Bierut<sup>40</sup>, Enda M. Byrne<sup>41,42</sup>, Jouke-Jan Hottenga<sup>43</sup>, Daniel L. Koller<sup>44</sup>, Massimo Mangino<sup>4</sup>, Tune H. Pers<sup>12,13,45,46</sup>, Laura M. Yerges-Armstrong<sup>47</sup>, Jing Hua Zhao<sup>41</sup>, Irene L. Andrusis<sup>48,49</sup>, Hoda Anton-Culver<sup>50</sup>, Femke Atsma<sup>51</sup>, Stefania Bandinelli<sup>52,53</sup>, Matthias W. Beckmann<sup>54</sup>, Javier Benitez<sup>55,56</sup>, Carl Blomqvist<sup>57</sup>, Stig E. Bojesen<sup>58,59</sup>, Manjeet K. Bolla<sup>6</sup>, Bernardo Bonanni<sup>60</sup>, Hiltrud Brauch<sup>61,62</sup>, Hermann Brenner<sup>63,64</sup>, Julie E. Buring<sup>3,10</sup>, Jenny Chang-Claude<sup>65</sup>, Stephen Chanock<sup>66</sup>, Jinhui Chen<sup>67,68</sup>, Georgia Chenevix-Trench<sup>69</sup>, J. Margriet Collée<sup>70</sup>, Fergus J. Couch<sup>71</sup>, David Couper<sup>72</sup>, Andrea D. Coviello<sup>73</sup>, Angela Cox<sup>74</sup>, Kamila Czene<sup>22</sup>, Adamo Pao D'adamio<sup>38,75</sup>, George Davey Smith<sup>26,27</sup>, Immaculata De Vivo<sup>76,77</sup>, Ellen W. Demerath<sup>78</sup>, Joe Dennis<sup>5</sup>, Peter Devilee<sup>79</sup>, Aida K. Dieffenbach<sup>63,64</sup>, Alison M. Dunning<sup>80</sup>, Gudny Eiriksdottir<sup>31</sup>, Johan G. Eriksson<sup>81,82,83,84</sup>, Peter A. Fasching<sup>54</sup>, Luigi Ferrucci<sup>85</sup>, Dieter Flesch-Janys<sup>86</sup>, Henrik Flyger<sup>87</sup>, Tatiana Foroud<sup>44</sup>, Lude Franke<sup>88</sup>, Melissa E. Garcia<sup>89</sup>, Montserrat Garcia-Closas<sup>90,91</sup>, Frank Geller<sup>20</sup>, Eco E. J. de Geus<sup>43,92</sup>, Graham G. Giles<sup>93,94</sup>, Daniel F. Gudbjartsson<sup>5,95</sup>, Vilundur Gudnason<sup>31,32</sup>, Pascal Guénel<sup>96,97</sup>, Suqin Guo<sup>98</sup>, Peter Hall<sup>22</sup>, Ute Hamann<sup>99</sup>, Robin Haring<sup>100</sup>, Catharina A. Hartman<sup>101</sup>, Andrew C. Heath<sup>102</sup>, Albert Hofman<sup>103</sup>, Maartje J. Hooning<sup>104</sup>, John L. Hopper<sup>94</sup>, Frank B. Hu<sup>76,77,105</sup>, David J. Hunter<sup>13,76,77</sup>, David Karasik<sup>10,106</sup>, Douglas P. Kiel<sup>106,107</sup>, Julia A. Knight<sup>108,109</sup>, Veli-Matti Kosma<sup>110,111</sup>, Zoltan Kutalik<sup>17,18</sup>, Sandra Lai<sup>29</sup>, Dieter Lambrechts<sup>112,113</sup>, Annika Lindblom<sup>114</sup>, Reedik Mägi<sup>11</sup>, Patrik K. Magnusson<sup>22</sup>, Arto Mannermaa<sup>110,111</sup>, Nicholas G. Martin<sup>61</sup>, Gisli Masson<sup>5</sup>, Patrick F. McArdle<sup>47</sup>, Wendy L. McGar<sup>127</sup>, Mads Melbye<sup>20,115</sup>, Kyriaki Michailidou<sup>6</sup>, Evelin Mihailov<sup>11,36</sup>, Lili Milani<sup>11</sup>, Roger L. Milne<sup>93,94</sup>, Heli Nevanlinna<sup>116</sup>, Patrick Neven<sup>117</sup>, Ellen A. Nohr<sup>118</sup>, Albertine J. Oldehinkel<sup>119</sup>, Ben A. Oostra<sup>39</sup>, Aarno Palotie<sup>19,120,121,122</sup>, Munro Peacock<sup>123</sup>, Nancy L. Pedersen<sup>22</sup>, Paolo Peterlongo<sup>124</sup>, Julian Peto<sup>125</sup>, Paul D. P. Pharoah<sup>80</sup>, Dirkje S. Postma<sup>126</sup>, Anneli Pouta<sup>81,127</sup>, Katri Pylkäs<sup>128</sup>, Paolo Radice<sup>129</sup>, Susan Ring<sup>26,27</sup>, Fernando Rivadeneira<sup>33,34,103</sup>, Antonietta Robino<sup>38,75</sup>, Lynda M. Rose<sup>9</sup>, Anja Rudolph<sup>65</sup>, Veikko Salomaa<sup>81</sup>, Serena Sanna<sup>29</sup>, David Schlessinger<sup>130</sup>, Marjanka K. Schmidt<sup>131</sup>, Melissa C. Southey<sup>132</sup>, Ulla Sovio<sup>133,134</sup>, Meir J. Stampfer<sup>76,77,105</sup>, Doris Stöckl<sup>135,136</sup>, Anna M. Storniolo<sup>123</sup>, Nicholas J. Timpson<sup>125,127</sup>, Jonathan Tyrer<sup>80</sup>, Jenny A. Visser<sup>33</sup>, Peter Vollenweider<sup>137</sup>, Henry Völzke<sup>138,139</sup>, Gerard Waeber<sup>137</sup>, Melanie Waldenberger<sup>140</sup>, Henri Wallaschowski<sup>100,139</sup>, Qin Wang<sup>6</sup>, Gonke Willemsen<sup>43</sup>, Robert Winqvist<sup>128</sup>, Bruce H. R. Wolffenbuttel<sup>141</sup>, Margaret J. Wright<sup>142</sup>, Australian Ovarian Cancer Study†, The GENICA Network†, kConFab†, The LifeLines Cohort

Study†, The InterAct Consortium†, Early Growth Genetics (EGG) Consortium†, Dorret I. Boomsma<sup>43</sup>, Michael J. Econs<sup>44,123</sup>, Kay-Tee Khaw<sup>45</sup>, Ruth J. F. Loos<sup>1,144</sup>, Mark I. McCarthy<sup>3,145,146</sup>, Grant W. Montgomery<sup>142</sup>, John P. Rice<sup>40</sup>, Elizabeth A. Steeten<sup>47,147</sup>, Unnur Thorsteinsdottir<sup>5,95</sup>, Cornelia M. van Duijn<sup>34,39,148</sup>, Behrooz Z. Alizadeh<sup>28</sup>, Sven Bergmann<sup>17,18</sup>, Eric Boerwinkle<sup>149</sup>, Heather A. Boyd<sup>20</sup>, Laura Crispin<sup>29</sup>, Paolo Gasparini<sup>38,75</sup>, Christian Gieger<sup>15</sup>, Tamara B. Harris<sup>89</sup>, Erik Ingelsson<sup>150</sup>, Marjo-Riitta Järvelin<sup>133,151,152,153,154</sup>, Peter Kraft<sup>76,155</sup>, Debbie Lawlor<sup>26,27</sup>, Andres Metspalu<sup>11,36</sup>, Craig E. Pennell<sup>16</sup>, Paul M. Ridker<sup>9,10</sup>, Harold Snieder<sup>28</sup>, Thorkild I. A. Sørensen<sup>156,157</sup>, Tim D. Spector<sup>4</sup>, David P. Strachan<sup>158</sup>, André G. Uitterlinden<sup>33,34,103</sup>, Nicholas J. Wareham<sup>1</sup>, Elisabeth Widen<sup>19</sup>, Marek Zygmunt<sup>159</sup>, Anna Murray<sup>7</sup>, Douglas F. Easton<sup>6</sup>, Kari Stefansson<sup>5,95\*</sup>, Joanne M. Murabito<sup>23,160\*</sup> & Ken K. Ong<sup>1,161\*</sup>

<sup>1</sup>MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Box 285 Institute of Metabolic Science, Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK. <sup>2</sup>University of Exeter Medical School, University of Exeter, Exeter EX1 2LU, UK. <sup>3</sup>Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford OX3 7BN, UK. <sup>4</sup>Department of Twin Research and Genetic Epidemiology, King's College London, London SE1 7EH, UK. <sup>5</sup>deCODE Genetics, Reykjavik IS-101, Iceland. <sup>6</sup>Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK. <sup>7</sup>Department of Epidemiology, Indiana University Richard M Fairbanks School of Public Health, Indianapolis, Indiana 46202, USA. <sup>8</sup>Indiana University Melvin and Bren Simon Cancer Center, Indianapolis, Indiana 46202, USA. <sup>9</sup>Division of Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts 02215, USA. <sup>10</sup>Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>11</sup>Estonian Genome Center, University of Tartu, Tartu, 51010, Estonia. <sup>12</sup>Divisions of Endocrinology and Genetics and Center for Basic and Translational Obesity Research, Boston Children's Hospital, Boston, Massachusetts 02115, USA. <sup>13</sup>Broad Institute of the Massachusetts Institute of Technology and Harvard University, 140 Cambridge, Massachusetts 02142, USA. <sup>14</sup>Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>15</sup>Institute of Genetic Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, D-85764 Neuherberg, Germany. <sup>16</sup>School of Women's and Infants' Health, The University of Western Australia, WA-6009, Australia. <sup>17</sup>Department of Medical Genetics, University of Lausanne, CH-1005 Lausanne, Switzerland. <sup>18</sup>Swiss Institute of Bioinformatics, CH-1015 Lausanne, Switzerland. <sup>19</sup>Institute for Molecular Medicine Finland (FIMM), University of Helsinki, FI-00014, Finland. <sup>20</sup>Department of Epidemiology Research, Statens Serum Institut, DK-2300 Copenhagen, Denmark. <sup>21</sup>Department of Epidemiology, University of North Carolina, Chapel Hill, North Carolina 27599-7400, USA. <sup>22</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, 17177 Stockholm, Sweden. <sup>23</sup>NHLBI's and Boston University's Framingham Heart Study, Framingham, Massachusetts 01702-5827, USA. <sup>24</sup>Science for Life Laboratory, Karolinska Institutet, Stockholm, Box 1031, 17121 Solna, Sweden. <sup>25</sup>Boston University School of Public Health, Department of Biostatistics, Boston, Massachusetts 02118, USA. <sup>26</sup>MRC Integrative Epidemiology Unit, University of Bristol, Bristol BS8 2BN, UK. <sup>27</sup>School of Social and Community Medicine, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK. <sup>28</sup>Department of Epidemiology, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, The Netherlands. <sup>29</sup>Institute of Genetics and Biomedical Research, National Research Council, Cagliari, 09042 Sardinia, Italy. <sup>30</sup>University of Sassari, Department of Biomedical Sciences, 07100 Sassari, Italy. <sup>31</sup>Icelandic Heart Association, IS-201 Kopavogur, Iceland. <sup>32</sup>University of Iceland, IS-101 Reykjavik, Iceland. <sup>33</sup>Department of Internal Medicine, Erasmus MC, 3015 GE Rotterdam, the Netherlands. <sup>34</sup>Netherlands Consortium on Health Aging and National Genomics Initiative, 2300 RC Leiden, the Netherlands. <sup>35</sup>Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, D-17475 Greifswald, Germany. <sup>36</sup>Department of Biotechnology, University of Tartu, 51010 Tartu, Estonia. <sup>37</sup>Hjelt Institute, University of Helsinki, FI-00014, Finland. <sup>38</sup>Institute for Maternal and Child Health - IRCCS "Burlo Garofolo", 34137 Trieste, Italy. <sup>39</sup>Genetic Epidemiology Unit Department of Epidemiology, Erasmus MC, 3015 GE, Rotterdam, the Netherlands. <sup>40</sup>Department of Psychiatry, Washington University, St Louis, Missouri 63110, USA. <sup>41</sup>The University of Queensland, Queensland Brain Institute, St Lucia, Queensland 4072, Australia. <sup>42</sup>QIMR Berghofer Medical Research Institute, Brisbane, Queensland 4006, Australia. <sup>43</sup>Department of Biological Psychology, VU University Amsterdam, van der Boechorststraat 1, 1081 BT, Amsterdam, The Netherlands. <sup>44</sup>Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana 46202-3082, USA. <sup>45</sup>Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts 02142, USA. <sup>46</sup>Center for Biological Sequence Analysis, Department of Systems Biology, Technical 142 University of Denmark, DK-2800 Lyngby, Denmark. <sup>47</sup>Program in Personalized and Genomic Medicine, and Department of Medicine, Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, Baltimore, Maryland 21201, USA. <sup>48</sup>Ontario Cancer Genetics Network, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Ontario M5G 1X5, Canada. <sup>49</sup>Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada. <sup>50</sup>Department of Epidemiology, University of California Irvine, Irvine, California 92697-7550, USA. <sup>51</sup>Sanquin Research, 6525 GA Nijmegen, The Netherlands. <sup>52</sup>Tuscany Regional Health Agency, Florence, Italy, I.O.T. and Department of Medical and Surgical Critical Care, University of Florence, 50134 Florence, Italy. <sup>53</sup>Geriatric Unit, Azienda Sanitaria di Firenze, 50122 Florence, Italy. <sup>54</sup>University Breast Center Franconia, Department of Gynecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, D-91054 Erlangen, Germany. <sup>55</sup>Human Genetics Group, Human Cancer Genetics Program, Spanish National Cancer Research Centre (CNIO), E-28029 Madrid, Spain. <sup>56</sup>Centro de Investigación en Red de Enfermedades Raras (CIBERER), E-46010 Valencia, Spain. <sup>57</sup>Department of Oncology, University of Helsinki and Helsinki University Central Hospital, FI-00100 Helsinki, Finland. <sup>58</sup>Copenhagen General Population Study, Herlev Hospital, Copenhagen University Hospital, University of Copenhagen, DK-2100 Copenhagen, Denmark. <sup>59</sup>Department of Clinical Biochemistry, Herlev Hospital,

Copenhagen University Hospital, University of Copenhagen, DK-2100 Copenhagen, Denmark. <sup>60</sup>Division of Cancer Prevention and Genetics, Istituto Europeo di Oncologia (IEO), 20139 Milan, Italy. <sup>61</sup>DrMargarete Fischer-Bosch-Institute of Clinical Pharmacology, D-70376 Stuttgart, Germany. <sup>62</sup>University of Tübingen, D-72074 Tübingen, Germany. <sup>63</sup>Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), D-69120 Heidelberg, Germany. <sup>64</sup>German Cancer Consortium (DKTK), D-69120 Heidelberg, Germany. <sup>65</sup>Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), D-69120 Heidelberg, Germany. <sup>66</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland 20892, USA. <sup>67</sup>Departments of Anatomy and Neurological Surgery, Indiana University school of Medicine, Indianapolis, Indiana 46202, USA. <sup>68</sup>Stark Neuroscience Research Center, Indiana University school of Medicine, Indianapolis, Indiana 46202, USA. <sup>69</sup>Department of Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Queensland 4006 Australia. <sup>70</sup>Department of Clinical Genetics, Erasmus University Medical Center, 3000 CA Rotterdam, The Netherlands. <sup>71</sup>Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota 55905, USA. <sup>72</sup>Department of Biostatistics, University of North Carolina, Chapel Hill, North Carolina 27599-7420, USA. <sup>73</sup>Boston University School of Medicine, Department of Medicine, Sections of Preventive Medicine and Endocrinology, Boston, Massachusetts 02118, USA. <sup>74</sup>Sheffield Cancer Research Centre, Department of Oncology, University of Sheffield, Sheffield S10 2RX, UK. <sup>75</sup>Department of Clinical Medical Sciences, Surgical and Health, University of Trieste, 34149 Trieste, Italy. <sup>76</sup>Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115, USA. <sup>77</sup>Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>78</sup>Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota 55455, USA. <sup>79</sup>Department of Human Genetics & Department of Pathology, Leiden University Medical Center, 2300 RC Leiden, The Netherlands. <sup>80</sup>Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge CB1 8RN, UK. <sup>81</sup>National Institute for Health and Welfare, P.O. Box 30, FI-00271 Helsinki, Finland. <sup>82</sup>Department of General Practice and Primary health Care, University of Helsinki, FI-00014 Helsinki, Finland. <sup>83</sup>Helsinki University Central Hospital, Unit of General Practice, FI-00029 HUS Helsinki, Finland. <sup>84</sup>Folkhalsan Research Centre, FI-00290 Helsinki, Finland. <sup>85</sup>Longitudinal Studies Section, Clinical Research Branch, Gerontology Research Center, National Institute on Aging, Baltimore, Maryland 20892, USA. <sup>86</sup>Department of Cancer Epidemiology/Clinical Cancer Registry and Institute for Medical Biometrics and Epidemiology, University Clinic Hamburg-Eppendorf, D-20246 Hamburg, Germany. <sup>87</sup>Department of Breast Surgery, Herlev Hospital, Copenhagen University Hospital, DK-2100 Copenhagen, Denmark. <sup>88</sup>Department of Genetics, University of Groningen, University Medical Centre Groningen, P.O. Box 72, 9700 AB Groningen, The Netherlands. <sup>89</sup>National Institute on Aging, National Institutes of Health, Baltimore, Maryland 20892, USA. <sup>90</sup>Division of Genetics and Epidemiology, Institute of Cancer Research, Sutton, Surrey SM2 5NG, UK. <sup>91</sup>Breakthrough Breast Cancer Research Centre, Division of Breast Cancer Research, The Institute of Cancer Research, London SW3 6JB, UK. <sup>92</sup>EMGO + Institute for Health and Care Research, VU University Medical Center, Van der Boechorststraat 7, 1081 BT, Amsterdam, The Netherlands. <sup>93</sup>Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Victoria 3004, Australia. <sup>94</sup>Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria 3010, Australia. <sup>95</sup>Faculty of Medicine, University of Iceland, IS-101 Reykjavik, Iceland. <sup>96</sup>Inserm (National Institute of Health and Medical Research), CESP (Center for Research in Epidemiology and Population Health), U1018, Environmental Epidemiology of Cancer, F-94807 Villejuif, France. <sup>97</sup>University Paris-Sud, UMRs 1018, F-94807 Villejuif, France. <sup>98</sup>Department of Obstetrics and Gynecology, Southern Medical University, 510515 Guangzhou, China. <sup>99</sup>Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), D-69120 Heidelberg, Germany. <sup>100</sup>Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, D-17475 Greifswald, Germany. <sup>101</sup>Department of Psychiatry, University of Groningen, University Medical Center Groningen, P.O. Box 72, 9700 AB Groningen, The Netherlands. <sup>102</sup>Washington University, Department of Psychiatry, St Louis, Missouri 63110, USA. <sup>103</sup>Department of Epidemiology, Erasmus MC, PO Box 2040, 3000 CA Rotterdam, the Netherlands. <sup>104</sup>Department of Medical Oncology, Erasmus University Medical Center, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands. <sup>105</sup>Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts 02115, USA. <sup>106</sup>Hebrew SeniorLife Institute for Aging Research, Boston, Massachusetts 02131, USA. <sup>107</sup>Department of Medicine, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>108</sup>Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Ontario M5G 1X5, Canada. <sup>109</sup>Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario M5T 3M7, Canada. <sup>110</sup>School of Medicine, Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, P.O. Box 1627, FI-70211 Kuopio, Finland. <sup>111</sup>Imaging Center, Department of Clinical Pathology, Kuopio University Hospital, P.O. Box 100, FI-70029 Kuopio, Finland. <sup>112</sup>Vesalius Research Center (VRC), VIB, 3000 Leuven, Belgium. <sup>113</sup>Laboratory for Translational Genetics, Department of Oncology, University of Leuven, 3000 Leuven, Belgium. <sup>114</sup>Department of Molecular Medicine and Surgery, Karolinska Institutet, SE-171 77 Stockholm, Sweden. <sup>115</sup>Department of Medicine, Stanford School of Medicine, Stanford, California 94305-5101, USA. <sup>116</sup>Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Central Hospital, P.O. Box 100, FI-00029 HUS Helsinki, Finland. <sup>117</sup>KULeuven (University of Leuven), Department of Oncology, Multidisciplinary Breast Center, University Hospitals Leuven, 3000 Leuven, Belgium. <sup>118</sup>Research Unit of Obstetrics & Gynecology, Institute of Clinical Research, University of Southern Denmark, DK-5000 Odense C, Denmark. <sup>119</sup>Interdisciplinary Center Psychopathology and Emotion Regulation, University of Groningen, University Medical Center Groningen, P.O. Box 30.001, 9700 RB Groningen, The Netherlands. <sup>120</sup>Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. <sup>121</sup>Program in Medical and Population Genetics, Broad Institute, Cambridge, Massachusetts 02142, USA. <sup>122</sup>Psychiatric & Neurodevelopmental Genetics

Unit, Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. <sup>123</sup>Department of Medicine, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA. <sup>124</sup>FOM, Fondazione Istituto FIRC di Oncologia Molecolare, 20139 Milan, Italy. <sup>125</sup>Non-communicable Disease Epidemiology Department, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK. <sup>126</sup>University Groningen, University Medical Center Groningen, Department Pulmonary Medicine and Tuberculosis, GRIAC Research Institute, P.O. Box 30.001, NL-9700 RB Groningen, The Netherlands. <sup>127</sup>Department of Obstetrics and Gynecology, Oulu University Hospital, P.O. Box 10, FI-90029 OYS Oulu, Finland. <sup>128</sup>Laboratory of Cancer Genetics and Tumor Biology, Department of Clinical Chemistry and Biocenter Oulu, University of Oulu, Oulu University Hospital/NordLab Oulu, P.O. Box 3000, FI-90014 Oulu, Finland. <sup>129</sup>Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori (INT), 20133 Milan, Italy. <sup>130</sup>National Institute on Aging, Intramural Research Program, Baltimore, Maryland 21224-6825, USA. <sup>131</sup>Netherlands Cancer Institute, Antoni van Leeuwenhoek hospital, Postbus 90203, 1006 BE Amsterdam, The Netherlands. <sup>132</sup>Department of Pathology, The University of Melbourne, Melbourne, Victoria 3010, Australia. <sup>133</sup>Department of Epidemiology and Biostatistics, MRC Health Protection Agency (HPA) Centre for Environment and Health, School of Public Health, Imperial College London, London W2 1PG, UK. <sup>134</sup>Department of Obstetrics and Gynaecology, University of Cambridge, Cambridge CB2 0SW, UK. <sup>135</sup>Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, D-8576 Neuherberg, Germany. <sup>136</sup>Department of Obstetrics and Gynaecology, Campus Grosshadern, Ludwig-Maximilians-University, D-81377 Munich, Germany. <sup>137</sup>Department of Internal Medicine, Lausanne University Hospital, CH-1015 Lausanne, Switzerland. <sup>138</sup>Institute for Community Medicine, University Medicine Greifswald, D-17475 Greifswald, Germany. <sup>139</sup>DZHK (German Centre for Cardiovascular Research), partner site Greifswald, D-17475 Greifswald, Germany. <sup>140</sup>Research Unit of Molecular Epidemiology, Helmholtz Zentrum München - German Research Center for Environmental Health, D-8576 Neuherberg, Germany. <sup>141</sup>Department of Endocrinology, University of Groningen, University Medical Centre Groningen, P.O. Box 72, 9700 AB Groningen, The Netherlands. <sup>142</sup>Queensland Institute of Medical Research, Brisbane, Queensland 4029, Australia. <sup>143</sup>Department of Public Health and Primary Care, Institute of Public Health, University of Cambridge, Cambridge CB2 0QQ, UK. <sup>144</sup>Genetics of Obesity and Related Metabolic Traits Program, The Charles Bronfman Institute for Personalized Medicine, The Mindich Child Health and Development Institute, Department of Preventive Medicine, Icahn School of Medicine at Mount Sinai, 1 Gustave L Levy Place, Box 1003, New York, New York 10029, USA. <sup>145</sup>NIHR Oxford Biomedical Research Centre, Churchill Hospital, Oxford OX3 7LE, UK. <sup>146</sup>Oxford Centre for Diabetes, Endocrinology, & Metabolism, University of Oxford, Churchill Hospital, Oxford OX3 7LJ, UK. <sup>147</sup>Geriatric Research and Education Clinical Center (GRECC) - Veterans Administration Medical Center, Baltimore, Maryland 21201, USA. <sup>148</sup>Centre of Medical Systems Biology, PO Box 9600, 2300 RC Leiden, the Netherlands. <sup>149</sup>Human Genetics Center and Divof Epidemiology, University of Houston, P.O. Box 20186, Texas 77025 USA. <sup>150</sup>Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Box 256, 751 05 Uppsala, Sweden. <sup>151</sup>Institute of Health Sciences, University of Oulu, P.O. Box 5000, FI-90014 Oulu, Finland. <sup>152</sup>Biocenter Oulu, University of Oulu, P.O. Box 5000, Aapistie 5A, FI-90014 Oulu, Finland. <sup>153</sup>Department of Children and Young People and Families, National Institute for Health and Welfare, Aapistie 1, Box 310, FI-90101 Oulu, Finland. <sup>154</sup>Unit of Primary Care, Oulu University Hospital, Kajaanintie 50, P.O. Box 20, FI-90220 Oulu, 90029 OYS, Finland. <sup>155</sup>Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts 02115, USA. <sup>156</sup>Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, DK-2200, Denmark. <sup>157</sup>Institute of Preventive Medicine, Bispebjerg and Frederiksberg Hospitals, The Capital Region, Copenhagen, DK-2000 Frederiksberg, Denmark. <sup>158</sup>Division of Population Health Sciences and Education, St George's, University of London, Cranmer Terrace, London SW17 0RE, UK. <sup>159</sup>Department of Obstetrics and Gynecology, University Medicine Greifswald, D-17475 Greifswald, Germany. <sup>160</sup>Boston University School of Medicine, Department of Medicine, Section of General Internal Medicine, Boston, Massachusetts 02118, USA. <sup>161</sup>Department of Paediatrics, University of Cambridge, Cambridge CB2 0QQ, UK.

\*These authors contributed equally to this work.

†Lists of participants and their affiliations appear in Supplementary Information.



## METHODS

**GWAS meta-analysis.** We performed an expanded GWAS meta-analysis for self-reported age at menarche in up to 182,416 women of European descent from 58 studies (Supplementary Table 1). All participants provided written informed consent and the studies were approved by the respective Local Research Ethics committees or Institutional Review Boards. Consistent with our previous analysis protocol<sup>4</sup>, women who reported their age at menarche as <9 years or >17 years were excluded from the analysis; birth year was included as the only covariate to allow for the secular trends in menarche timing. Genome-wide SNP array data were available on up to 132,989 women from 57 studies. Each study imputed genotype data based on HapMap Phase II CEU build 35 or 36. Data on an additional 49,427 women from the Breast Cancer Association Consortium (BCAC) were generated on the Illumina iSelect “iCOGS” array<sup>31</sup>. This array included up to ~25,000 SNPs, or their proxy markers, that showed sub-genome-wide associations ( $P < 0.0022$ ) with age at menarche in our earlier GWAS<sup>4</sup>. SNPs were excluded from individual study data sets if they were poorly imputed or were rare (minor allele frequency < 1%). Test statistics for each study were adjusted using study-specific genomic control inflation factors and where appropriate individual studies performed additional adjustments for relatedness (Supplementary Table 1). Association statistics for each of the 2,441,815 autosomal SNPs that passed QC in at least half of the studies were combined across studies in a fixed effects inverse-variance meta-analysis implemented in METAL<sup>32</sup>.

On meta-analysis, 3,915 SNPs reached the genome-wide significance threshold ( $P < 5 \times 10^{-8}$ ) for association with age at menarche (Fig. 1). The overall GC inflation factor was 1.266, consistent with an expected high yield of true positive findings in large-scale GWAS meta-analysis of highly polygenic traits<sup>33</sup>.

**Selection of independent signals.** Given the genome-wide results of the meta-analysis, SNPs showing evidence for association at genome-wide significant  $P$ -values were selected and clumped based on a physical (kb) threshold <1 megabase. The lead SNPs of the 105 clumps formed constitute the list of SNPs independently associated with age at menarche (Extended Data Tables 1–4).

To augment this list we performed approximate conditional analysis using GCTA software<sup>34</sup>, where the LD between variants was estimated from the Northern Finland Birth Cohort (NFBC66) consisting of 5,402 individuals of European ancestry with GWAS data imputed using CEU haplotypes from Hapmap Phase II. Assuming that the LD correlations between SNPs more than 10 Mb away or on different chromosomes are zero, we performed the GCTA model selection to select SNPs independently associated with age at menarche at genome-wide significant  $P$ -values. This software selected as independently associated with age at menarche 115 SNPs at 98 loci, 11 of which had two or more signals of association (six loci contained two signals, four loci contained three signals, and one locus contained four signals). Plots of all 106 loci are available at <http://www.reprogen.org>. SNPs with A/T or C/G alleles were excluded from this analysis to prevent strand issues leading to false-positive results.

To summarize the information obtained from the single-SNP and GCTA analyses, the 105 SNPs selected from the uni-variate analysis and the 115 SNPs selected from the GCTA model selection analysis were combined into a single list of signals independently associated with age at menarche (Supplementary Table 2), using the following selection process (Extended Data Fig. 1). For loci with no evidence of allelic heterogeneity, if the uni-variate signal was genome-wide significant, the lead uni-variate SNP was selected (94 independent association signals follow this criterion); otherwise the lead GCTA SNP was selected instead (one independent signal). For loci where evidence for allelic heterogeneity was found, all signals identified in the GCTA joint model were selected if GCTA selected the uni-variate index SNP (21 independent signals at 8 loci) or a very good proxy ( $r^2 > 0.8$ ) (7 independent signals at 3 loci). When instead GCTA selected a SNP independent from the uni-variate index SNP, both the lead uni-variate SNP and all signals identified in the GCTA joint model were selected (0 independent signals).

To determine likely causal genes at each locus, we used a combination of criteria. The gene nearest to each top SNP was selected by default. This gene was replaced or added to if the top SNP was (in high LD with) an expression quantitative-trait locus (eQTL) or a non-synonymous variant in another gene, or if there was an alternative neighbouring biological candidate gene. 31/123 signals mapped as eQTLs in data from Westra *et al.* (E)<sup>10</sup>, five were annotated as non-synonymous functional (F), 60 as biological candidates (C), and four mapped to gene deserts (nearest gene > 500 kb) (Supplementary Tables 6–8). We also used publicly available whole blood and adipose tissue methylation-QTL data to map 9/123 signals to *cis*-acting changes in methylation level (Extended Data Table 5)<sup>9</sup>.

**Follow up in the EPIC-InterAct study.** We used an independent sample of 8689 women from the EPIC-InterAct study<sup>35</sup> to follow up our menarche signals. To test associations between each identified SNP and age at menarche with correction for cryptic relatedness, we ran a linear mixed model association test implemented in GCTA<sup>34</sup> (–mlma-loco option), adjusting for birth year, disease status and research centre. Given the relatively small sample size compared to our discovery set, directional consistency with results from the discovery-meta analysis was assessed using

a binomial sign test. Variance explained by menarche loci was estimated using restricted maximum likelihood analysis in GCTA<sup>34</sup>. In addition to the 123 confirmed menarche loci, variance explained in subsets of menarche loci below the genome-wide significance thresholds was also assessed.

**eQTL analyses.** In order to estimate the potential downstream regulatory effects of age at menarche associated variants, we used publicly available blood eQTL data (downloadable from <http://genenetwork.nl/bloodeqtlbrowser/>) from a recently published paper by Westra *et al.*<sup>10</sup>. Westra *et al.* conducted *cis*-eQTL mapping by testing, for a large set of genes, all SNPs (HapMap2 panel) within 250 kb of the transcription start site of the gene for association with total RNA expression level of the gene. The publicly available data contain, for each gene, a list of all SNPs that were found to be significantly associated with gene expression using a false discovery rate (FDR) of 5%. For a detailed description of the quality control measures applied to the original data, see Westra *et al.*<sup>10</sup>. Their meta-analysis was based on a pooled sample of 5,311 individuals from 7 population-based cohorts with gene expression levels measured from full blood. We used the software tool SNAP (<http://www.broadinstitute.org/mpg/snap/>) to identify variants in close linkage disequilibrium ( $r^2 \geq 0.8$ ) with the trait associated variants. All eQTL effects at FDR 5% and also lists of the strongest SNP effect for all the significant genes are shown in Supplementary Table 7.

Index SNPs (or highly correlated proxies) were also interrogated against a collected database of eQTL results from a range of tissues. Blood cell related eQTL studies included fresh lymphocytes<sup>36</sup>, fresh leukocytes<sup>37</sup>, leukocyte samples in individuals with Crohn's disease<sup>38</sup>, whole blood samples<sup>39–43</sup>, lymphoblastoid cell lines (LCL) derived from asthmatic children<sup>44,45</sup>, HapMap LCL from 3 populations<sup>46</sup>, a separate study on HapMap CEU LCL<sup>47</sup>, additional LCL population samples<sup>48–50</sup> (and Mangravite *et al.* (unpublished)), CD19<sup>+</sup> B cells<sup>51</sup>, primary PHA-stimulated T cells<sup>48</sup>, CD4<sup>+</sup> T cells<sup>52</sup>, peripheral blood monocytes<sup>51,53,54</sup>, CD11<sup>+</sup> dendritic cells before and after *Mycobacterium tuberculosis* infection<sup>55</sup>, Micro-RNA QTLs<sup>56</sup> and DNase-I QTLs<sup>57</sup> were also queried for LCL. Non-blood cell tissue eQTLs searched included omental and subcutaneous adipose<sup>39,50,58</sup>, stomach<sup>58</sup>, endometrial carcinomas<sup>59</sup>, ER+ and ER– breast cancer tumour cells<sup>60</sup>, brain cortex<sup>53,61,62</sup>, pre-frontal cortex<sup>63,64</sup>, frontal cortex<sup>65</sup>, temporal cortex<sup>62,65</sup>, pons<sup>65</sup>, cerebellum<sup>62,65</sup>, 3 additional large studies of brain regions including prefrontal cortex, visual cortex and cerebellum, respectively<sup>66</sup>, liver<sup>58,67–70</sup>, osteoblasts<sup>71</sup>, intestine<sup>72</sup>, lung<sup>73</sup>, skin<sup>50,74</sup> and primary fibroblasts<sup>48</sup>. Micro-RNA QTLs were also queried for gluteal and abdominal adipose<sup>75</sup>. Only results that reach study-wise significance thresholds in their respective data sets were included (Supplementary Table 6). Expression data was also available on adipose tissue and whole blood samples from deCODE where parent-of-origin-specific analyses were possible.

**Parent-of-origin-specific associations.** Evidence for parent-of-origin-specific allelic associations at imprinted loci was sought in the deCODE Study, which included 35,377 women with parental origins of alleles determined by a combination of genealogy and long-range phasing as previously described<sup>6</sup>. Briefly, using SNP chip data in each proband, genome-wide, long range phasing was applied to overlapping tiles, each 6 centimorgan (cM) in length, with 3 cM overlap between consecutive tiles. For each tile, the parental origins of the two phased haplotypes were determined regardless of whether the parents of the proband were chip-typed. Using the Icelandic genealogy database, for each of the two haplotypes of a proband, a search was performed to identify, among those individuals also known to carry the same haplotype, the closest relative on each of the paternal and maternal sides. Results for the two haplotypes were combined into a robust single-tile score reflecting the relative likelihood of the two possible parental origin assignments. Haplotypes from consecutive tiles were then stitched together based on sharing at the overlapping region. For haplotypes derived by stitching, a contig-score for parental origin was computed by summing the individual single-tile scores. Similarly, parent-of-origin-specific allelic associations at imprinted loci were also sought in the deCODE blood cells and adipose tissue expression data sets.

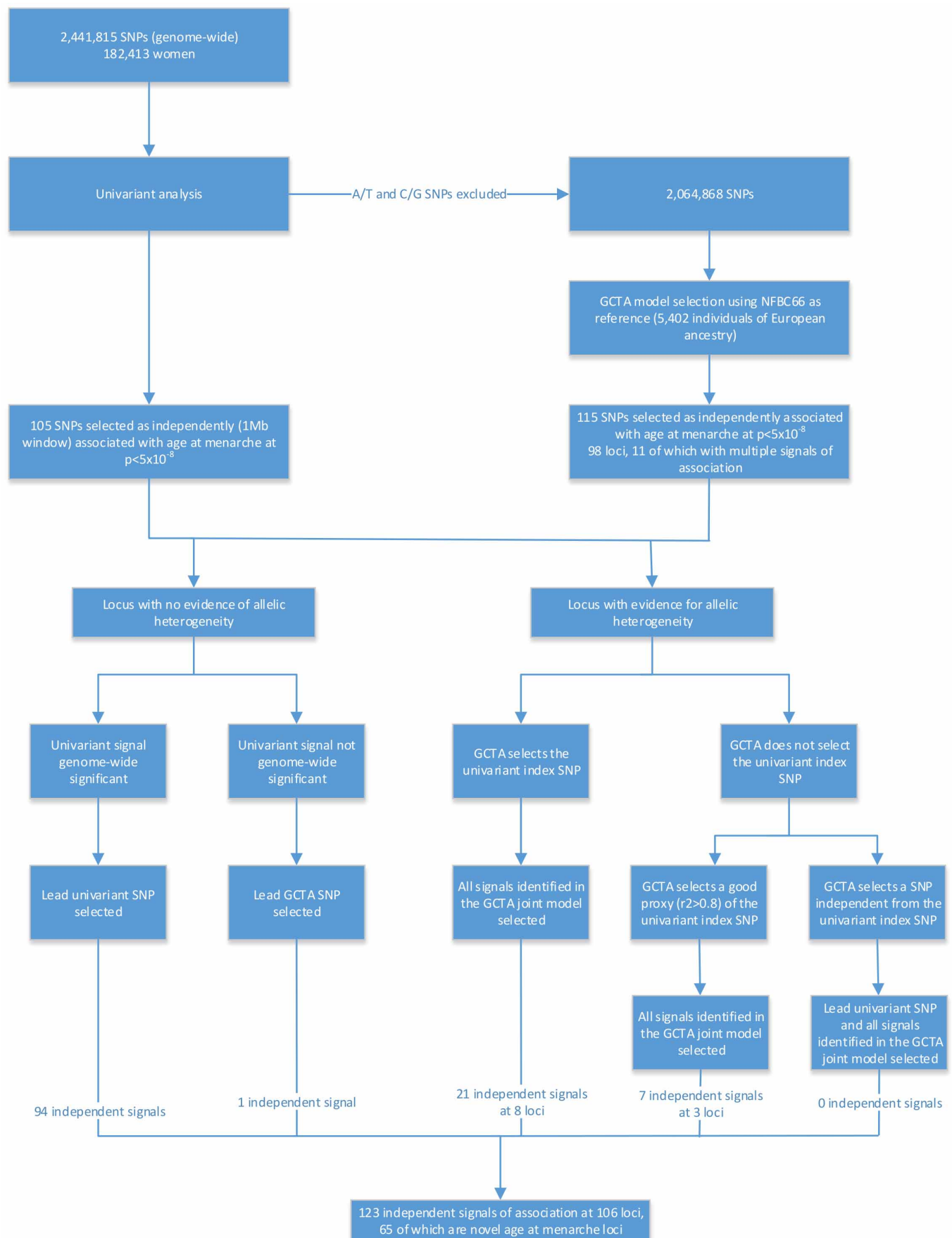
**Pathway analyses.** Meta-Analysis Gene-set Enrichment of variaNT Associations (MAGENTA) was used to explore pathway-based associations in the full GWAS data set. MAGENTA implements a gene set enrichment analysis (GSEA) based approach, as previously described<sup>76</sup>. Briefly, each gene in the genome is mapped to a single index SNP with the lowest  $P$ -value within a 110 kb upstream, 40 kb downstream window. This  $P$ -value, representing a gene score, is then corrected for confounding factors such as gene size, SNP density and LD-related properties in a regression model. Genes within the HLA-region were excluded from analysis due to difficulties in accounting for gene density and LD patterns. Each mapped gene in the genome is then ranked by its adjusted gene score. At a given significance threshold (95th and 75th percentiles of all gene scores), the observed number of gene scores in a given pathway, with a ranked score above the specified threshold percentile, is calculated. This observed statistic is then compared to 1,000,000 randomly permuted pathways of identical size. This generates an empirical GSEA  $P$ -value for each pathway. Significance was determined when an individual pathway reached a false discovery rate (FDR) < 0.05 in either analysis. In total, 2529 pathways from Gene Ontology, PANTHER, KEGG



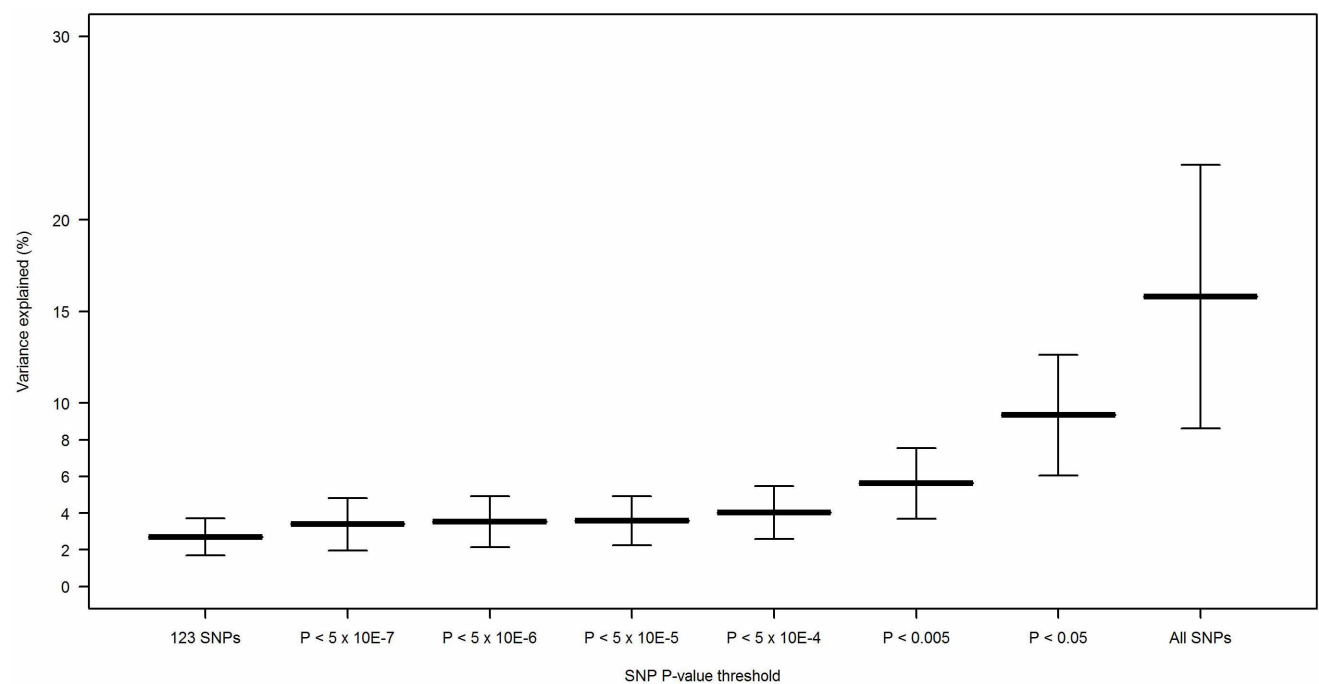
and Ingenuity were tested for enrichment of multiple modest associations with age at menarche. MAGENTA software was also used for enrichment testing of custom gene sets.

**Relevance of menarche loci to other traits.** We assessed the relevance of identified menarche loci to other traits by comparing SNPs significantly associated with age at menarche with published GWAS findings or by using publicly available data from the Genetic Investigation of Anthropometric Traits (GIANT) consortium<sup>21,22</sup> and the GEnetic Factors for OS (GEFOS) consortium<sup>27</sup>. In addition, we requested look-ups up the 123 menarche SNPs for association with puberty timing assessed by Tanner staging in the Early Growth Genetics (EGG) consortium<sup>77</sup>.

31. Michailidou, K. *et al.* Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nature Genet.* **45**, 353–361 (2013).
32. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genome-wide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
33. Yang, J. *et al.* Genomic inflation factors under polygenic inheritance. *Eur. J. Hum. Genet.* **19**, 807–812 (2011).
34. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
35. The InterAct Consortium Design and cohort description of the InterAct Project: an examination of the interaction of genetic and lifestyle factors on the incidence of type 2 diabetes in the EPIC Study. *Diabetologia* **54**, 2272–2282 (2011).
36. Göring, H. H. H. *et al.* Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nature Genet.* **39**, 1208–1216 (2007).
37. Idaghdour, Y. *et al.* Geographical genomics of human leukocyte gene expression variation in southern Morocco. *Nature Genet.* **42**, 62–67 (2010).
38. Heap, G. A. *et al.* Complex nature of SNP genotype effects on gene expression in primary human leukocytes. *BMC Med. Genomics* **2**, 1 (2009).
39. Emilsson, V. *et al.* Genetics of gene expression and its effect on disease. *Nature* **452**, 423–428 (2008).
40. Fehrmann, R. S. N. *et al.* Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. *PLoS Genet.* **7**, e1002197 (2011).
41. Mehta, D. *et al.* Impact of common regulatory single-nucleotide variants on gene expression profiles in whole blood. *Eur. J. Hum. Genet.* **21**, 48–54 (2013).
42. Maeda, T. *et al.* The correlation between clinical laboratory data and telomeric status of male patients with metabolic disorders and no clinical history of vascular events. *Aging Male* **14**, 21–26 (2011).
43. Sasayama, D. *et al.* Identification of single nucleotide polymorphisms regulating peripheral blood mRNA expression with genome-wide significance: an eQTL study in the Japanese population. *PLoS ONE* **8**, e54967 (2013).
44. Dixon, A. L. *et al.* A genome-wide association study of global gene expression. *Nature Genet.* **39**, 1202–1207 (2007).
45. Liang, L. *et al.* A cross-platform analysis of 14,177 expression quantitative trait loci derived from lymphoblastoid cell lines. *Genome Res.* **23**, 716–726 (2013).
46. Stranger, B. E. *et al.* Population genomics of human gene expression. *Nature Genet.* **39**, 1217–1224 (2007).
47. Kwan, T. *et al.* Genome-wide analysis of transcript isoform variation in humans. *Nature Genet.* **40**, 225–231 (2008).
48. Dimas, A. S. *et al.* Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* **325**, 1246–1250 (2009).
49. Cusanovich, D. A. *et al.* The combination of a genome-wide association study of lymphocyte count and analysis of gene expression data reveals novel asthma candidate genes. *Hum. Mol. Genet.* **21**, 2111–2123 (2012).
50. Grundberg, E. *et al.* Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nature Genet.* **44**, 1084–1089 (2012).
51. Fairfax, B. P. *et al.* Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nature Genet.* **44**, 502–510 (2012).
52. Murphy, A. *et al.* Mapping of numerous disease-associated expression polymorphisms in primary peripheral blood CD4+ lymphocytes. *Hum. Mol. Genet.* **19**, 4745–4757 (2010).
53. Heinzen, E. L. *et al.* Tissue-specific genetic control of splicing: implications for the study of complex traits. *PLoS Biol.* **6**, e1 (2008).
54. Zeller, T. *et al.* Genetics and beyond—the transcriptome of human monocytes and disease susceptibility. *PLoS ONE* **5**, e10693 (2010).
55. Barreiro, L. B. *et al.* Deciphering the genetic architecture of variation in the immune response to *Mycobacterium tuberculosis* infection. *Proc. Natl Acad. Sci. USA* **109**, 1204–1209 (2012).
56. Huang, R. S. *et al.* Population differences in microRNA expression and biological implications. *RNA Biol.* **8**, 692–701 (2011).
57. Degner, J. F. *et al.* DNase I sensitivity QTLs are a major determinant of human expression variation. *Nature* **482**, 390–394 (2012).
58. Greenawald, D. M. *et al.* A survey of the genetics of stomach, liver, and adipose gene expression from a morbidly obese cohort. *Genome Res.* **21**, 1008–1016 (2011).
59. Kompass, K. S. & Witte, J. S. Co-regulatory expression quantitative trait loci mapping: method and application to endometrial cancer. *BMC Med. Genomics* **4**, 6 (2011).
60. Li, Q. *et al.* Integrative eQTL-based analyses reveal the biology of breast cancer risk loci. *Cell* **152**, 633–641 (2013).
61. Webster, J. A. *et al.* Genetic control of human brain transcript expression in Alzheimer disease. *Am. J. Hum. Genet.* **84**, 445–458 (2009).
62. Zou, F. *et al.* Brain expression genome-wide association study (eGWAS) identifies human disease-associated variants. *PLoS Genet.* **8**, e1002707 (2012).
63. Colantuoni, C. *et al.* Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* **478**, 519–523 (2011).
64. Liu, C. *et al.* Whole-genome association mapping of gene expression in the human prefrontal cortex. *Mol. Psychiatry* **15**, 779–784 (2010).
65. Gibbs, J. R. *et al.* Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* **6**, e1000952 (2010).
66. Zhang, B. *et al.* Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell* **153**, 707–720 (2013).
67. Schadt, E. E. *et al.* Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* **6**, e107 (2008).
68. Innocenti, F. *et al.* Identification, replication, and functional fine-mapping of expression quantitative trait loci in primary human liver tissue. *PLoS Genet.* **7**, e1002078 (2011).
69. Sulzbacher, S., Schroeder, I. S., Truong, T. T. & Wobus, A. M. Activin A-induced differentiation of embryonic stem cells into endoderm and pancreatic progenitors—the influence of differentiation factors and culture conditions. *Stem Cell Rev.* **5**, 159–173 (2009).
70. Schröder, A. *et al.* Genomics of ADME gene expression: mapping expression quantitative trait loci relevant for absorption, distribution, metabolism and excretion of drugs in human liver. *Pharmacogenomics J.* **13**, 12–20 (2013).
71. Grundberg, E. *et al.* Population genomics in a disease targeted primary cell model. *Genome Res.* **19**, 1942–1952 (2009).
72. Kabackchiev, B. & Silverberg, M. S. Expression quantitative trait loci analysis identifies associations between genotype and gene expression in human intestine. *Gastroenterology* **144**, 1488–1496e3 (2013).
73. Hao, K. *et al.* Lung eQTLs to help reveal the molecular underpinnings of asthma. *PLoS Genet.* **8**, e1003029 (2012).
74. Ding, J. *et al.* Gene expression in skin and lymphoblastoid cells: refined statistical method reveals extensive overlap in cis-eQTL signals. *Am. J. Hum. Genet.* **87**, 779–789 (2010).
75. Rantalainen, M. *et al.* MicroRNA expression in abdominal and gluteal adipose tissue is associated with mRNA expression levels and partly genetically driven. *PLoS ONE* **6**, e27338 (2011).
76. Segrè, A. V., Groop, L., Mootha, V. K., Daly, M. J. & Altshuler, D. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet.* **6**, e1001058 (2010).
77. Cousminer, D. L. *et al.* Genome-wide association study of sexual maturation in males and females highlights a role for body mass and menarche loci in male puberty. *Hum. Mol. Genet.*; Epub ahead of print (2014).



Extended Data Figure 1 | Flow chart illustrating the selection criteria used to identify independent signals for age at menarche.



**Extended Data Figure 2 | Estimates of genetic variance explained.** Variance explained by combined sets of SNPs defined by their strength of association in the EPIC-InterAct replication sample ( $N = 8,689$ ) in the discovery set.

Extended Data Table 1 | Details of the 123 independent signals for menarche timing at 106 genomic loci—signals no. 1 to 30

Locus	SNP	Location <sup>1</sup>	Novel (r-sq) <sup>2</sup>	N	Alleles / Freq <sup>3</sup>	Uni-variate Model <sup>4</sup>		Joint Model <sup>5</sup>		Gene <sup>6</sup>
						Beta (se)	P	Beta (se)	P	
1	rs2274465	1-43894144	Yes	179348	c/g/0.66	0.03 (0.005)	1.7E-09	n/a	n/a	<i>KDM4A</i> <sup>[NC]</sup> , <i>PTPRF</i> <sup>[EC]</sup>
2	rs10789181	1-65589155	Yes	177560	a/g/0.39	0.03 (0.005)	3.5E-08	n/a	n/a	<i>LEPR</i> <sup>[C]</sup>
3	rs3101336	1-72523773	Yes	182404	t/c/0.4	0.04 (0.005)	5.2E-13	n/a	n/a	<i>NEGR1</i> <sup>[NC]</sup>
4	rs7514705	1-74779308	Yes	179631	c/t/0.56	0.04 (0.005)	1.8E-16	n/a	n/a	<i>TNNI3K</i> <sup>[NJ]</sup> , <i>TYW3</i> <sup>[E]</sup>
5	rs11165924	1-98148036	Yes	174006	a/g/0.69	0.03 (0.006)	2.2E-09	n/a	n/a	<i>DPYD</i> <sup>[NJ]</sup>
6	rs11578152	1-102349609	Yes	179433	g/a/0.44	0.03 (0.005)	4.5E-08	n/a	n/a	<i>OLFM3</i> <sup>[NJ]</sup>
7	rs466639	1-163661506	No (Same)	179432	c/t/0.87	0.08 (0.007)	2.4E-24	n/a	n/a	<i>RXRG</i> <sup>[NC]</sup>
8	rs543874	1-176156103	No (0.91)	179613	a/g/0.8	0.05 (0.006)	1.4E-15	n/a	n/a	<i>SEC16B</i> <sup>[NJ]</sup>
9	rs6427782	1-198064962	Yes	175785	a/g/0.51	0.03 (0.005)	4.6E-08	n/a	n/a	<i>NR5A2</i> <sup>[NC]</sup>
10	rs951366	1-203951975	Yes	179567	t/c/0.6	0.03 (0.005)	1.7E-08	n/a	n/a	<i>NUCKS1</i> <sup>[NE]</sup> , <i>RAB7L1</i> <sup>[E]</sup>
11	rs2947411	2-604168	No (Same)	179608	a/g/0.17	0.06 (0.007)	1.8E-19	n/a	n/a	<i>TMEM18</i> <sup>[NC]</sup>
12	rs6747380	2-56441253	No (1.0)	182377	a/g/0.17	0.07 (0.007)	5.6E-28	n/a	n/a	<i>CCDC85A</i> <sup>[NJ]</sup>
13	rs268067	2-59734549	Yes	179406	a/g/0.8	0.04 (0.006)	3.3E-08	n/a	n/a	<i>BCL11A</i> <sup>[N~800kb]</sup>
14	rs6758290	2-105231258	Yes	167496	t/c/0.5	0.04 (0.005)	6.6E-13	n/a	n/a	<i>GPR45</i> <sup>[NJ]</sup>
15	rs12472911	2-141944979	No (Same)	182269	c/t/0.2	0.04 (0.006)	6.7E-10	n/a	n/a	<i>LRP1B</i> <sup>[NJ]</sup>
16a	rs17236969	2-156460705	Yes	162496	t/c/0.14	0.05 (0.008)	2.6E-09	0.05 (0.008)	1.0E-08	<i>NR4A2</i> <sup>[NC]</sup>
16b	rs4369815	2-156835210	No (Same)	174922	t/g/0.93	0.06 (0.01)	1.5E-10	0.06 (0.01)	5.5E-10	<i>NR4A2</i> <sup>[NC]</sup>
17a	rs1400974	2-199346935	No (0.78)	179605	a/g/0.64	0.05 (0.005)	8.3E-20	0.04 (0.005)	3.0E-17	<i>SATB2</i> <sup>[NJ]</sup>
17b	rs17233066	2-199352283	No (0.22)	168273	c/t/0.93	0.09 (0.014)	6.1E-11	0.08 (0.014)	1.8E-09	<i>SATB2</i> <sup>[NJ]</sup>
17c	rs17266097	2-199983454	Yes	179181	t/c/0.42	0.04 (0.005)	3.3E-18	0.04 (0.005)	2.4E-16	<i>SATB2</i> <sup>[NJ]</sup>
18	rs6770162	3-24686017	Yes	179304	a/g/0.51	0.04 (0.005)	1.5E-12	n/a	n/a	<i>THRB</i> <sup>[NC]</sup>
19a	rs7647973	3-49485935	No (0.74)	179667	a/g/0.26	0.05 (0.006)	1.3E-16	0.05 (0.006)	2.4E-16	<i>WDR6</i> <sup>[EC]</sup> , <i>UBA7</i> <sup>[C]</sup>
19b	rs6762477	3-50068213	No (Same)	138679	g/a/0.44	0.04 (0.006)	7.8E-12	0.04 (0.006)	2.2E-11	<i>WDR6</i> <sup>[EC]</sup> , <i>UBA7</i> <sup>[C]</sup>
20	rs7642134	3-86999572	No (Same)	182263	g/a/0.61	0.04 (0.005)	3.0E-16	n/a	n/a	<i>POU1F1</i> <sup>[C]</sup> ( <i>PIT1</i> )
21	rs9849248	3-88323964	Yes	179654	c/t/0.15	0.04 (0.007)	1.9E-08	n/a	n/a	<i>ZNF654</i> <sup>[NEF]</sup> , <i>HTR1F</i> <sup>[C]</sup>
22	rs11715566	3-119045126	No (0.97)	179637	t/c/0.5	0.05 (0.005)	2.4E-27	n/a	n/a	<i>IGSF11</i> <sup>[N~1Mb]</sup>
23	rs2687729	3-129377916	No (Same)	179617	g/a/0.27	0.04 (0.006)	1.0E-10	n/a	n/a	<i>EEFSEC</i> <sup>[NE]</sup>
24	rs2600959	3-134098154	No (0.97)	174583	a/g/0.34	0.04 (0.005)	4.1E-11	n/a	n/a	<i>ACAD11</i> <sup>[E]</sup>
25	rs13067731	3-138472681	Yes	179330	t/c/0.16	0.04 (0.007)	1.0E-09	n/a	n/a	<i>IL20RB</i> <sup>[NJ]</sup>
26	rs900400	3-158281469	Yes	179649	t/c/0.61	0.03 (0.005)	2.3E-11	n/a	n/a	<i>LEKR1</i> <sup>[NE]</sup> , <i>CCNL1</i> <sup>[C]</sup>
27	rs939317	3-185528493	No (0.8)	179622	g/a/0.74	0.04 (0.006)	3.0E-12	n/a	n/a	<i>EIF4G1</i> <sup>[NJ]</sup>
28	rs16860328	3-187118379	No (0.93)	179646	g/a/0.42	0.04 (0.005)	1.4E-16	n/a	n/a	<i>TRA2B</i> <sup>[NJ]</sup> , <i>IGF2BP2</i> <sup>[C]</sup>
29	rs1038903	4-28361152	Yes	179610	t/c/0.73	0.04 (0.006)	2.0E-11	n/a	n/a	<i>PCDH7</i> <sup>[N~2Mb]</sup>
30	rs10938397	4-44877284	Yes	179167	a/g/0.57	0.04 (0.005)	4.0E-13	n/a	n/a	<i>GNPDA2</i> <sup>[NJ]</sup>

<sup>1</sup>All positions mapped to Hapmap build 36.<sup>2</sup>Novel indicates previously unidentified loci. If the locus was established, r-sq refers to the linkage disequilibrium between the reported SNP and the previous signal. Some regions with known associations and no prior evidence for allelic heterogeneity now have multiple independent signals.<sup>3</sup>Alleles/freq refers to the menarche age-increasing allele (from the uni-variate SNP discovery), and the decreasing allele/increasing allele frequencies from meta-analysis study estimates.<sup>4</sup>Uni-variate models included only one SNP per model.<sup>5</sup>Joint models were performed using GCTA software. These models approximate conditional analysis; that is, the effect estimates are adjusted for the effects of other neighbouring SNPs.<sup>6</sup>Gene refers to the consensus gene(s) reported at that locus mapped using 4 approaches: N, nearest; C, biological candidate; F, 1000 Genomes missense variant in high LD ( $r^2 > 0.8$ ); E, gene expression linked by eQTL. See Supplementary Tables 5, 7 and 8 for more information.



Extended Data Table 2 | Details of the 123 independent signals for menarche timing at 106 genomic loci—signals no. 31 to 58

Locus	SNP	Location <sup>1</sup>	Novel (r-sq) <sup>2</sup>	N	Alleles / Freq <sup>3</sup>	Uni-variate Model <sup>4</sup>		Joint Model <sup>5</sup>		Gene <sup>6</sup>
						Beta (se)	P	Beta (se)	P	
31	rs13135934	4-95426711	Yes	178661	c/g/0.4	0.03 (0.005)	1.1E-10	n/a	n/a	SMARCD1 <sup>[NEF]</sup>
32	rs3733631	4-104860552	Yes	179623	c/g/0.15	0.05 (0.007)	4.8E-13	n/a	n/a	TACR3 <sup>[NC]</sup>
33	rs1532331	5-43152587	Yes	179201	g/t/0.32	0.03 (0.005)	3.5E-09	n/a	n/a	ZNF131 <sup>[NEC]</sup> , GHR <sup>[C]</sup>
34	rs17086188	5-95871610	Yes	176967	a/g/0.94	0.07 (0.013)	3.6E-08	n/a	n/a	PCSK1 <sup>[NC]</sup>
35	rs2511130	5-110887696	Yes	179429	g/a/0.73	0.04 (0.006)	2.8E-10	n/a	n/a	STARD4 <sup>[NEC]</sup>
36	rs13179411	5-133928412	No (0.53)	179579	t/g/0.17	0.06 (0.007)	3.4E-20	n/a	n/a	PHF15 <sup>[N]</sup> , TCF7 <sup>[E]</sup>
37	rs17171818	5-137752902	No (1.0)	182224	c/t/0.77	0.04 (0.006)	8.9E-14	n/a	n/a	KDM3B <sup>[NC]</sup> , BRD8 <sup>[C]</sup>
38	rs7701886	5-153527602	Yes	179664	a/g/0.58	0.03 (0.005)	4.5E-08	n/a	n/a	GALNT10 <sup>[N]</sup>
39	rs9647570	5-167302841	Yes	179600	g/t/0.14	0.05 (0.007)	1.4E-11	n/a	n/a	TENM2 <sup>[NC]</sup>
40	rs6555855	5-168682315	Yes	179462	g/a/0.23	0.04 (0.006)	2.4E-09	n/a	n/a	SLIT3 <sup>[N]</sup>
41	rs16896742	6-30030719	Yes	171665	g/a/0.38	0.04 (0.006)	3.2E-10	n/a	n/a	HLA-A <sup>[N]</sup>
42	rs2479724	6-41998960	Yes	179630	t/c/0.45	0.03 (0.005)	1.2E-12	n/a	n/a	BYSL <sup>[NE]</sup> , FRS3 <sup>[C]</sup>
43	rs988913	6-54864267	Yes	182407	c/t/0.66	0.04 (0.005)	1.4E-12	n/a	n/a	FAM83B <sup>[N]</sup> , HCRT2 <sup>[C]</sup>
44	rs9475752	6-56888700	Yes	178646	c/t/0.81	0.04 (0.006)	8.3E-12	n/a	n/a	DST <sup>[N]</sup> , BEND6 <sup>[E]</sup>
45	rs9447700	6-77224806	Yes	179648	c/t/0.69	0.03 (0.005)	5.6E-09	n/a	n/a	IMPG1 <sup>[N]</sup>
46a	rs9321659	6-100222813	Yes	182356	a/g/0.13	0.06 (0.008)	2.5E-16	0.06 (0.008)	2.9E-16	SIM1 <sup>[C]</sup> , MCHR2 <sup>[C]</sup>
46b	rs4840086	6-100315159	No (Same)	179666	a/g/0.58	0.04 (0.005)	9.2E-14	0.04 (0.005)	4.3E-13	SIM1 <sup>[C]</sup> , MCHR2 <sup>[C]</sup>
46c	rs13196561	6-100866891	Yes	182278	c/a/0.78	0.04 (0.006)	8.4E-12	0.06 (0.006)	3.4E-20	SIM1 <sup>[NC]</sup> , MCHR2 <sup>[C]</sup>
46d	rs239198	6-101240798	Yes	179496	t/c/0.46	0.03 (0.005)	2.5E-08	0.04 (0.005)	3.1E-15	SIM1 <sup>[C]</sup> , ASCC3 <sup>[NEF]</sup>
47a	rs4946632	6-105207901	Yes	132973	c/t/0.1	0.01 (0.01)	0.14	-0.07 (0.01)	3.1E-12	LIN28B <sup>[C]</sup>
47b	rs2153127	6-105455237	Yes	182110	t/c/0.52	0.08 (0.005)	5.5E-59	0.03 (0.006)	2.1E-09	LIN28B <sup>[EC]</sup>
47c	rs7759938	6-105485647	No (Same)	179557	c/t/0.32	0.12 (0.005)	7.8E-110	0.11 (0.006)	1.2E-69	LIN28B <sup>[NC]</sup>
48	rs4895808	6-126823127	No (1.0)	179655	c/t/0.54	0.03 (0.005)	4.8E-13	n/a	n/a	CENPW <sup>[NE]</sup> , NCOA7 <sup>[C]</sup>
49	rs6938574	6-128432673	Yes	178428	t/c/0.16	0.04 (0.007)	2.4E-09	n/a	n/a	PTPRK <sup>[N]</sup>
50	rs6933660	6-151845447	Yes	182379	c/a/0.69	0.03 (0.005)	1.3E-09	n/a	n/a	ESR1 <sup>[C]</sup>
51	rs1079866	7-41436618	No (Same)	172036	g/c/0.15	0.07 (0.007)	9.3E-24	n/a	n/a	INHBA <sup>[NC]</sup>
52	rs6964833	7-73739845	Yes	171484	t/c/0.75	0.04 (0.006)	5.3E-12	n/a	n/a	GTF2I <sup>[NC]</sup>
53	rs11767400	7-121947978	Yes	179658	a/c/0.3	0.04 (0.006)	4.1E-11	n/a	n/a	CADPS2 <sup>[N]</sup>
54a	rs2688325	8-3754618	Yes	182244	t/c/0.29	0.03 (0.006)	2.1E-09	0.03 (0.006)	9.7E-10	CSMD1 <sup>[N]</sup>
54b	rs7828501	8-4547489	Yes	179434	g/a/0.45	0.04 (0.005)	1.2E-13	0.04 (0.005)	2.8E-15	CSMD1 <sup>[N]</sup>
54c	rs7463166	8-4821198	Yes	179542	a/g/0.63	0.03 (0.005)	1.3E-08	0.03 (0.005)	5.9E-09	CSMD1 <sup>[N]</sup>
55	rs16918254	8-53931766	Yes	179635	a/g/0.92	0.05 (0.009)	1.4E-08	n/a	n/a	NPBWR1 <sup>[NC]</sup>
56	rs7821178	8-78256392	No (Same)	179533	c/a/0.65	0.04 (0.005)	7.3E-17	n/a	n/a	PEX2 <sup>[N]</sup>
57	rs1469039	8-140720961	Yes	174755	a/g/0.19	0.05 (0.007)	3.5E-12	n/a	n/a	KCNK9 <sup>[N]</sup>
58	rs4875053	8-144944399	Yes	136628	g/c/0.44	0.03 (0.006)	1.3E-08	n/a	n/a	SCRIB <sup>[N]</sup> , PARP10 <sup>[E]</sup>

<sup>1</sup>All positions mapped to Hapmap build 36.<sup>2</sup>Novel indicates previously unidentified loci. If the locus was established, r-sq refers to the linkage disequilibrium between the reported SNP and the previous signal. Some regions with known associations and no prior evidence for allelic heterogeneity now have multiple independent signals.<sup>3</sup>Alleles/freq refers to the menarche age-increasing allele (from the uni-variate SNP discovery), and the decreasing allele/increasing allele frequencies from meta-analysis study estimates.<sup>4</sup>Uni-variate models included only one SNP per model.<sup>5</sup>Joint models were performed using GCTA software. These models approximate conditional analysis; that is, the effect estimates are adjusted for the effects of other neighbouring SNPs.<sup>6</sup>Gene refers to the consensus gene(s) reported at that locus mapped using 4 approaches: N, nearest; C, biological candidate; F, 1000 Genomes missense variant in high LD ( $r^2 > 0.8$ ); E, gene expression linked by eQTL. See Supplementary Tables 5, 7 and 8 for more information.

Extended Data Table 3 | Details of the 123 independent signals for menarche timing at 106 genomic loci—signals no. 59 to 87

Locus	SNP	Location <sup>1</sup>	Novel (r-sq) <sup>2</sup>	N	Alleles / Freq <sup>3</sup>	Uni-variate Model <sup>4</sup>		Joint Model <sup>5</sup>		Gene <sup>6</sup>
						Beta (se)	P	Beta (se)	P	
59a	rs7037266	9-6932940	Yes	179488	a/c/0.37	0.03 (0.005)	4.7E-09	0.03 (0.005)	3.5E-09	<i>KDM4C</i> <sup>[NC]</sup>
59b	rs913588	9-7164673	Yes	182403	g/a/0.49	0.03 (0.005)	5.8E-11	0.03 (0.005)	3.8E-11	<i>KDM4C</i> <sup>[NEC]</sup>
60	rs7865468	9-10264080	Yes	179418	a/g/0.7	0.03 (0.005)	1.3E-07	0.03 (0.005)	1.9E-08	<i>PTPRD</i> <sup>[NJ]</sup>
61	rs7853970	9-85905386	Yes	169702	t/c/0.47	0.03 (0.005)	2.3E-09	n/a	n/a	<i>RMI1</i> <sup>[NJ]</sup> , <i>NTRK2</i> <sup>[C]</sup>
62a	rs10816359	9-107797491	Yes	169277	t/g/0.86	0.04 (0.008)	1.6E-08	0.05 (0.008)	1.2E-12	<i>TMEM38B</i> <sup>[NJ]</sup>
62b	rs10453225	9-107960041	No (0.73)	179631	g/t/0.68	0.09 (0.005)	5.8E-66	0.07 (0.006)	3.5E-33	<i>TMEM38B</i> <sup>[NJ]</sup>
62c	rs10739221	9-108100651	No (0.42)	179624	c/t/0.77	0.08 (0.006)	3.9E-41	0.05 (0.007)	1.9E-11	<i>TMEM38B</i> <sup>[NJ]</sup>
63	rs11792861	9-110849116	Yes	179618	a/c/0.7	0.04 (0.005)	1.7E-11	n/a	n/a	<i>TMEM245</i> <sup>[NE]</sup>
64a	rs10980854	9-113090178	Yes	181999	a/g/0.06	0.06 (0.011)	1.3E-08	0.06 (0.011)	4.3E-09	<i>ZNF483</i> / <i>OR2K2</i> <sup>[NJ]</sup>
64b	rs10980921	9-113319733	No (0.12)	172160	c/t/0.09	0.09 (0.009)	1.7E-23	0.09 (0.009)	4.3E-23	<i>ZNF483</i> / <i>OR2K2</i> <sup>[NJ]</sup>
65	rs1874984	10-1721871	Yes	179112	c/g/0.47	0.04 (0.005)	1.9E-12	n/a	n/a	<i>ADARB2</i> <sup>[NJ]</sup>
66	rs12571664	10-121698919	Yes	179629	t/c/0.79	0.04 (0.006)	3.3E-10	n/a	n/a	<i>SEC23IP</i> <sup>[NE]</sup>
67	rs1915146	10-126836204	Yes	182401	g/a/0.4	0.03 (0.005)	3.7E-08	n/a	n/a	<i>CTBP2</i> <sup>[NC]</sup>
68	rs7104764	11-219977	Yes	179664	g/a/0.25	0.03 (0.006)	3.7E-08	n/a	n/a	<i>SIRT3</i> <sup>[NEC]</sup>
69	rs4929947	11-8596570	No (1.0)	179331	g/c/0.36	0.04 (0.005)	2.6E-12	n/a	n/a	<i>TRIM66</i> <sup>[NEF]</sup>
70	rs11022756	11-13272015	No (0.88)	179401	a/c/0.29	0.05 (0.006)	7.4E-20	n/a	n/a	<i>ARNTL</i> <sup>[NJ]</sup> , <i>PTH</i> <sup>[C]</sup>
71	rs7103411	11-27656701	Yes	179656	c/t/0.21	0.04 (0.006)	2.6E-11	n/a	n/a	<i>BDNF</i> <sup>[NC]</sup> , <i>LGR4</i> <sup>[C]</sup>
72	rs16918636	11-29080758	Yes	182237	t/c/0.79	0.03 (0.006)	3.2E-08	n/a	n/a	<i>FSHB</i> <sup>[CN~1Mb]</sup>
73	rs4756059	11-46107195	No (0.65)	179478	t/c/0.92	0.07 (0.01)	4.5E-13	n/a	n/a	<i>PHF21A</i> <sup>[NJ]</sup>
74	rs2063730	11-77726172	No (0.75)	179293	c/a/0.18	0.05 (0.007)	2.3E-12	n/a	n/a	<i>GAB2</i> <sup>[NJ]</sup> , <i>THRSP</i> <sup>[C]</sup>
75	rs10895140	11-100941931	Yes	179647	g/a/0.66	0.04 (0.005)	6.7E-14	n/a	n/a	<i>TRPC6</i> <sup>[NJ]</sup> , <i>PGR</i> <sup>[C]</sup>
76	rs11215400	11-114557845	Yes	179376	c/a/0.27	0.04 (0.006)	6.8E-11	n/a	n/a	<i>CADM1</i> <sup>[NJ]</sup>
77	rs1461503	11-122350285	No (0.34)	179603	c/a/0.57	0.05 (0.005)	2.7E-26	n/a	n/a	<i>BSX</i> <sup>[NC]</sup>
78	rs7955374	12-46166416	Yes	179419	t/c/0.13	0.04 (0.008)	9.5E-09	n/a	n/a	<i>VDR</i> <sup>[C]</sup>
79	rs7138803	12-48533735	Yes	174834	g/a/0.62	0.04 (0.005)	1.7E-12	n/a	n/a	<i>BCDIN3D</i> <sup>[NJ]</sup>
80	rs6563739	13-39137785	Yes	179667	g/t/0.34	0.03 (0.005)	2.3E-11	n/a	n/a	<i>COG6</i> <sup>[NE]</sup>
81	rs1324913	13-73533589	Yes	182393	g/t/0.65	0.03 (0.005)	3.1E-10	n/a	n/a	<i>KLF12</i> <sup>[NJ]</sup>
82	rs9560113	13-110981349	No (1.0)	179359	g/a/0.28	0.05 (0.006)	2.1E-17	n/a	n/a	<i>TEX29</i>
83	rs1254337	14-59990278	Yes	179658	t/a/0.31	0.04 (0.005)	2.1E-16	n/a	n/a	<i>SIX6</i> <sup>[NJ]</sup>
84	rs1958560	14-65106548	Yes	179655	a/g/0.59	0.03 (0.005)	3.7E-08	n/a	n/a	<i>FUT8</i> <sup>[NE]</sup>
85a	rs10144321	14-99952158	Yes	179595	a/g/0.75	0.04 (0.006)	9.0E-15	0.04 (0.006)	1.1E-14	<i>DLK1</i> <sup>[C]</sup> , <i>WDR25</i> <sup>[E]</sup>
85b	rs7141210	14-100252223	Yes	172034	t/c/0.34	0.03 (0.005)	5.8E-09	0.03 (0.005)	4.1E-09	<i>DLK1</i> <sup>[NEC]</sup>
86	rs12148769	15-21703187	Yes	182411	g/a/0.9	0.05 (0.008)	5.2E-11	n/a	n/a	<i>MKRN3</i> <sup>[C]</sup> , <i>MAGEL2</i> <sup>[C]</sup>
87	rs3743266	15-58568805	No (Same)	182389	t/c/0.68	0.04 (0.005)	2.4E-13	n/a	n/a	<i>RORA</i> <sup>[NC]</sup>

<sup>1</sup>All positions mapped to Hapmap build 36.<sup>2</sup>Novel indicates previously unidentified loci. If the locus was established, r-sq refers to the linkage disequilibrium between the reported SNP and the previous signal. Some regions with known associations and no prior evidence for allelic heterogeneity now have multiple independent signals.<sup>3</sup>Alleles/freq refers to the menarche age-increasing allele (from the uni-variate SNP discovery), and the decreasing allele/increasing allele frequencies from meta-analysis study estimates.<sup>4</sup>Uni-variate models included only one SNP per model.<sup>5</sup>Joint models were performed using GCTA software. These models approximate conditional analysis; that is, the effect estimates are adjusted for the effects of other neighbouring SNPs.<sup>6</sup>Gene refers to the consensus gene(s) reported at that locus mapped using 4 approaches: N, nearest; C, biological candidate; F, 1000 Genomes missense variant in high LD ( $r^2 > 0.8$ ); E, gene expression linked by eQTL. See Supplementary Tables 5, 7 and 8 for more information.

Extended Data Table 4 | Details of the 123 independent signals for menarche timing at 106 genomic loci—signals no. 88 to 106

Locus	SNP	Location <sup>1</sup>	Novel (r-sq) <sup>2</sup>	N	Alleles / Freq <sup>3</sup>	Uni-variate Model <sup>4</sup>		Joint Model <sup>5</sup>		Gene <sup>6</sup>
						Beta (se)	P	Beta (se)	P	
88	rs8032675	15-65746518	No (0.39)	179630	t/c/0.4	0.04 (0.005)	2.1E-13	n/a	n/a	MAP2K5 <sup>[NJ]</sup>
89	rs12915845	15-86843471	Yes	179535	c/t/0.58	0.03 (0.005)	2.7E-12	n/a	n/a	DET1 <sup>[NE]</sup>
90	rs246185	16-14302933	Yes (0.84)	177773	c/t/0.33	0.04 (0.006)	6.8E-16	n/a	n/a	MKL2 <sup>[NJ]</sup>
91	rs12446632	16-19842890	Yes	182401	a/g/0.13	0.04 (0.007)	1.3E-08	n/a	n/a	GPRC5B <sup>[NC]</sup>
92	rs1129700	16-29825535	Yes	181797	t/c/0.44	0.03 (0.005)	2.3E-09	n/a	n/a	KCTD13 <sup>[NJ]</sup> , TBX6 <sup>[EC]</sup>
93	rs8050136	16-52373776	No (1.0)	182365	c/a/0.6	0.04 (0.005)	1.7E-17	n/a	n/a	FTO <sup>[NC]</sup>
94a	rs1364063	16-68146073	No (Same)	182393	c/t/0.43	0.05 (0.005)	6.2E-21	0.04 (0.005)	4.8E-18	COG4 <sup>[C]</sup> , NFAT5 <sup>[NJ]</sup>
94b	rs929843	16-68603249	Yes	177329	a/c/0.23	0.04 (0.006)	1.2E-11	0.04 (0.006)	5.9E-09	COG4 <sup>[C]</sup> , WWP2 <sup>[NJ]</sup>
95	rs7215990	17-5975555	Yes	170053	g/a/0.76	0.04 (0.006)	1.9E-08	n/a	n/a	WSCD1 <sup>[NE]</sup> , ALOX15B <sup>[E]</sup>
96	rs9635759	17-46968784	No (Same)	179649	a/g/0.32	0.05 (0.005)	1.7E-24	n/a	n/a	CA10 <sup>[NJ]</sup>
97	rs244293	17-50585721	Yes	179560	g/a/0.6	0.03 (0.005)	4.2E-11	n/a	n/a	STXBPA <sup>[NE]</sup>
98	rs12607903	18-3807134	Yes	179171	c/t/0.3	0.04 (0.005)	5.4E-11	n/a	n/a	DLGAP1 <sup>[NJ]</sup>
99	rs2137289	18-43006123	No (0.74)	178617	a/g/0.59	0.05 (0.005)	8.2E-20	n/a	n/a	SKOR2 <sup>[NJ]</sup>
100	rs652260	19-7806562	Yes	182356	t/c/0.54	0.03 (0.005)	9.9E-09	n/a	n/a	EVI5L <sup>[NJ]</sup> , RETN <sup>[C]</sup>
101	rs889122	19-9856867	No (0.33)	179397	g/t/0.72	0.04 (0.006)	1.6E-13	n/a	n/a	OLFM2 <sup>[NJ]</sup> , RDH8 <sup>[C]</sup>
102	rs10423674	19-18678903	No (Same)	182377	a/c/0.34	0.04 (0.005)	9.2E-12	n/a	n/a	CRTC1 <sup>[NC]</sup>
103	rs852069	20-17070593	No (Same)	182413	g/a/0.64	0.04 (0.005)	1.2E-13	n/a	n/a	PCSK2 <sup>[NC]</sup>
104	rs2836950	21-39526299	Yes	178602	c/g/0.64	0.03 (0.005)	6.2E-11	n/a	n/a	BRWD1 <sup>[NC]</sup>
105	rs13053505	22-37575564	Yes	177596	g/t/0.8	0.04 (0.007)	3.0E-08	n/a	n/a	NPTXR <sup>[NE]</sup> , CBX7 <sup>[C]</sup>
106	rs6009583	22-48063650	Yes	181839	c/t/0.74	0.03 (0.006)	4.6E-08	n/a	n/a	C22orf34 <sup>[NJ]</sup>

<sup>1</sup>All positions mapped to Hapmap build 36.<sup>2</sup>Novel indicates previously unidentified loci. If the locus was established, r-sq refers to the linkage disequilibrium between the reported SNP and the previous signal. Some regions with known associations and no prior evidence for allelic heterogeneity now have multiple independent signals.<sup>3</sup>Alleles/freq refers to the menarche age-increasing allele (from the uni-variate SNP discovery), and the decreasing allele/increasing allele frequencies from meta-analysis study estimates.<sup>4</sup>Uni-variate models included only one SNP per model.<sup>5</sup>Joint models were performed using GCTA software. These models approximate conditional analysis; that is, the effect estimates are adjusted for the effects of other neighbouring SNPs.<sup>6</sup>Gene refers to the consensus gene(s) reported at that locus mapped using 4 approaches: N, nearest; C, biological candidate; F, 1000 Genomes missense variant in high LD ( $r^2 > 0.8$ ); E, gene expression linked by eQTL. See Supplementary Tables 5, 7 and 8 for more information.

Extended Data Table 5 | Methylation QTLs based on Illumina 450K whole blood and adipose methylome data in 648 twins

Locus	SNP	Consensus gene	Methylation probe <sup>1,2</sup>	Adipose tissue			Whole blood	
				Beta <sup>3</sup>	SE	P	Beta <sup>3</sup>	P
16b	rs4369815	<i>NR4A2</i> (N,C)	cg14912644	0.006	0.002	7.3E-04	-	-
33	rs1532331	<i>ZNF131</i> (N,E,C), <i>GHR</i> (C )	cg18254356	-0.01	0.003	4.4E-04	-	-
36	rs13179411	<i>PHF15</i> (N), <i>TCF7</i> (E )	cg00043364	-0.02	0.003	7.9E-11	-0.35	7.3E-03
64b	rs10980921	<i>ZNF483</i> / <i>OR2K2</i> (N)	cg01294431	0.01	0.002	1.1E-08	-	-
67	rs1915146	<i>CTBP2</i> (N,C)	cg17191109	0.01	0.001	6.9E-16	0.75	2.8E-18
83	rs1254337	<i>SIX6</i> (N)	cg00157572	-0.005	0.001	3.8E-05	-	-
85b	rs7141210	<i>DLK1</i> (N,E,C)	cg17008318	0.02	0.002	1.3E-18	-	-
100	rs652260	<i>EVI5L</i> (N), <i>RETN</i> (C )	cg06793867	-0.03	0.003	1.3E-23	-	-
100	rs652260	<i>EVI5L</i> (N), <i>RETN</i> (C )	cg14209047	0.01	0.002	2.4E-12	0.35	1.9E-04
100	rs652260	<i>EVI5L</i> (N), <i>RETN</i> (C )	cg15974673	-0.03	0.003	4.8E-27	-0.6	2.1E-11
102	rs10423674	<i>CRTC1</i> (N,C)	cg19861427	-0.007	0.002	1.4E-05	-	-

<sup>1</sup>Methylation-QTLs were derived for associations between genotypes and methylation in 648 adipose samples from the MuTHER study using a 1% FDR level, corresponding to  $P < 8.6 \times 10^{-41}$ . Significant methylation-QTLs were also tested for replication in whole blood in 200 individuals.

<sup>2</sup>Methylation data available from ref. 9.

<sup>3</sup>Methylation betas are presented per menarche-age-increasing allele.



Extended Data Table 6 | MAGENTA pathway analyses

Database	Gene set	Genes (mapped) <sup>1</sup>	95th percentile enrichment cut-off			75th percentile enrichment cut-off		
			P	FDR	Enrichment <sup>2</sup> Exp. (obs.)	P	FDR	Enrichment <sup>2</sup> Exp. (obs.)
Panther	GABA <sub>B</sub> receptor II signaling	9 (9)	8.00E-04	9.25E-03	0 (4)	9.70E-03	1.12E-01	2 (6)
Panther	Angiotensin II-stimulated signaling through G proteins and beta-arrestin	5 (5)	6.00E-04	1.39E-02	0 (3)	1.39E-02	9.78E-02	1 (4)
GOTERM	Regulation of transcription	991 (844)	1.30E-05	2.65E-01	42 (69)	1.00E-06	7.00E-04	211 (271)
GOTERM	Transcription factor activity	947 (788)	4.51E-03	4.19E-01	39 (55)	2.40E-05	3.89E-02	197 (242)
BIOCARTA	ETC_PATHWAY	12 (9)	3.78E-01	5.59E-01	0 (1)	1.20E-03	4.23E-02	2 (7)
GOTERM	Chromatin assembly or disassembly	38 (31)	4.69E-01	9.05E-01	2 (2)	1.10E-05	1.15E-02	8 (19)
Panther	5HT3 type receptor mediated signaling	7 (5)	1.00E+00	9.27E-01	0 (0)	1.10E-03	1.65E-02	1 (5)
Custom	Nuclear hormone receptors	57 (55)	6.00E-05	6.00E-05	3 (11)	4.58E-03	9.60E-03	14 (23)
Custom	Lysine specific demethylases	24 (24)	5.60E-03	5.60E-03	1 (5)	1.24E-01	1.24E-01	6 (9)
Custom	Mendelian pubertal disorders <sup>3</sup>	20 (18)	5.30E-02	5.30E-02	1 (3)	1.38E-01	1.38E-01	5 (7)

Results are shown for database pathways and custom pathways that reached study-wise statistical significance (FDR <0.05).

<sup>1</sup>Genes denotes number of genes in pathway (number of genes successfully mapped by MAGENTA).

<sup>2</sup>Enrichment denotes expected number of genes at enrichment threshold (observed number of genes).

<sup>3</sup>Genes for Mendelian pubertal disorders, as described in refs 2 and 3.

Extended Data Table 7 | GABA<sub>B</sub> receptor II signalling pathway genes

Gene	Gene P	Gene size(kb)	Number of SNPs	Number of Recombination Hotspots	Best SNP	Best SNP p value
<i>ADCY8</i>	2.87E-03	260	489	9	rs4392877	6.83E-08
<i>ADCY6</i>	4.89E-03	23	92	3	rs2446999	8.70E-07
<i>GABBR1</i>	9.32E-03	31	405	2	rs1362126	1.33E-06
<i>PRKAR2A</i>	9.04E-03	97	59	2	rs11713694	1.99E-06
<i>PRKAR2B</i>	2.81E-01	117	209	4	rs2244846	1.17E-03
<i>ADCY9</i>	3.42E-01	154	309	7	rs879150	1.51E-03
<i>GABBR2</i>	5.51E-01	421	698	10	rs2485144	2.86E-03
<i>ADCY1</i>	6.08E-01	149	184	3	rs10951832	1.27E-02
<i>ADCY5</i>	7.13E-01	164	207	5	rs9880405	2.31E-02

# The neurotrophic factor receptor RET drives haematopoietic stem cell survival and function

Diogo Fonseca-Pereira<sup>1\*</sup>, Sílvia Arroiz-Madeira<sup>1\*</sup>, Mariana Rodrigues-Campos<sup>1</sup>, Inês A. M. Barbosa<sup>1</sup>, Rita G. Domingues<sup>1</sup>, Teresa Bento<sup>1</sup>, Afonso R. M. Almeida<sup>1</sup>, Hélder Ribeiro<sup>1</sup>, Alexandre J. Potocnik<sup>2,3</sup>, Hideki Enomoto<sup>4,5</sup> & Henrique Veiga-Fernandes<sup>1</sup>

**Haematopoiesis is a developmental cascade that generates all blood cell lineages in health and disease. This process relies on quiescent haematopoietic stem cells capable of differentiating, self renewing and expanding upon physiological demand<sup>1,2</sup>. However, the mechanisms that regulate haematopoietic stem cell homeostasis and function remain largely unknown. Here we show that the neurotrophic factor receptor RET (rearranged during transfection) drives haematopoietic stem cell survival, expansion and function. We find that haematopoietic stem cells express RET and that its neurotrophic factor partners are produced in the haematopoietic stem cell environment. Ablation of *Ret* leads to impaired survival and reduced numbers of haematopoietic stem cells with normal differentiation potential, but loss of cell-autonomous stress response and reconstitution potential. Strikingly, RET signals provide haematopoietic stem cells with critical *Bcl2* and *Bcl2l1* surviving cues, downstream of p38 mitogen-activated protein (MAP) kinase and cyclic-AMP-response element binding protein (CREB) activation. Accordingly, enforced expression of RET downstream targets, *Bcl2* or *Bcl2l1*, is sufficient to restore the activity of *Ret* null progenitors *in vivo*. Activation of RET results in improved haematopoietic stem cell survival, expansion and *in vivo* transplantation efficiency. Remarkably, human cord-blood progenitor expansion and transplantation is also improved by neurotrophic factors, opening the way for exploration of RET agonists in human haematopoietic stem cell transplantation. Our work shows that neurotrophic factors are novel components of the haematopoietic stem cell microenvironment, revealing that haematopoietic stem cells and neurons are regulated by similar signals.**

Haematopoietic stem cells (HSCs) are mostly quiescent in adulthood but can become proliferative upon physiological demand<sup>1,2</sup>. Autonomic nerves have been shown to be in close proximity to HSCs, raising the possibility that both cell types might be regulated through similar mechanisms<sup>3–5</sup>. Neurotrophic factors are key to neuron function and include the glial cell-line derived neurotrophic factor (GDNF) family of ligands (GFLs), which signal through the RET tyrosine kinase receptor in neurons, kidney and lymphoid cell subsets<sup>6–8</sup>.

Initially we determined the expression of the canonical GFL receptor RET in fetal liver Lin<sup>–</sup>Sca1<sup>+</sup>cKit<sup>+</sup> (LSK) cells. When compared with myeloid progenitors (Lin<sup>–</sup>Sca1<sup>–</sup>cKit<sup>+</sup>), LSKs expressed high levels of *Ret* and its co-receptors *Gfra1*, *Gfra2* and *Gfra3* (Extended Data Fig. 1a). *Ret* expression was higher in Lin<sup>–</sup>Sca1<sup>+</sup>cKit<sup>+</sup>CD150<sup>+</sup>CD48<sup>–</sup> haematopoietic stem cells (HSCs), while multipotent progenitors (Lin<sup>–</sup>Sca1<sup>+</sup>cKit<sup>+</sup>CD150<sup>–</sup>CD48<sup>+</sup>) expressed this gene poorly (Fig. 1a)<sup>6,8–10</sup>. *Ret* expression by fetal HSCs was comparable to lymphoid tissue initiator cells (LTin), which are functionally dependent on RET (Fig. 1a, b)<sup>6,11</sup>, while bone marrow HSCs expressed low levels of *Ret* (Extended Data Fig. 1b, c). Interestingly, cells known to support HSCs expressed the RET ligands GDNF, neurturin (NRTN) and artemin (ARTN) (Fig. 1c

and Extended Data Fig. 1d, e)<sup>12,13</sup>. In agreement, the fetal liver and bone marrow HSC environment revealed the presence of GFLs in the vicinity of candidate HSCs, suggesting a role of RET in these cells (Fig. 1d, e and Extended Data Fig. 1f, g)<sup>2</sup>.

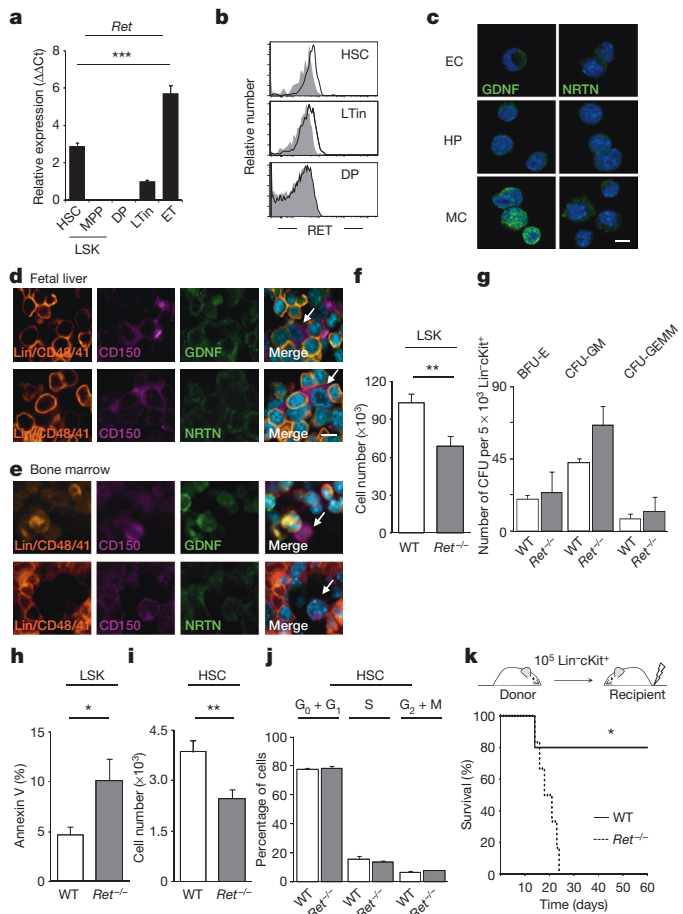
To test this hypothesis we analysed *Ret* null mice<sup>14</sup>. Embryonic day (E)14.5 *Ret*-deficient LSKs and fetal liver cellularity were reduced in *Ret*<sup>–/–</sup> embryos (Fig. 1f and Extended Data Fig. 2a). However, the differentiation potential of *Ret*<sup>–/–</sup> LSKs was intact as revealed by normal colony-forming units (CFU) and *ex vivo* differentiation (Fig. 1g and Extended Data Fig. 2b). Interestingly, *Ret* null LSKs were highly susceptible to apoptosis (Fig. 1h and Extended Data Fig. 2c) and *Ret* deficiency resulted in decreased HSC numbers with normal cell cycle profile (Fig. 1i, j and Extended Data Fig. 2d, e). These findings led us to investigate long-term HSC transplantation. Despite similar homing capacity (Extended Data Fig. 2f), *Ret*-deficient progenitors failed to rescue lethally irradiated mice (Fig. 1k).

To evaluate the fate of *Ret* null progenitors, we performed competitive transplantation assays. Fetal *Ret*<sup>–/–</sup> progenitors and wild type (WT) littermate controls were co-transplanted with equal numbers of third-party WT progenitors that ensured host survival (Fig. 2a). Analysis of recipient mice revealed that *Ret* null progenitors lost their transplantation fitness across all blood cell lineages (Fig. 2b and Extended Data Fig. 2g). Accordingly, bone marrow analysis 4 months after transplantation showed minute frequencies of *Ret*-deficient LSKs (Fig. 2c). Sequentially we performed highly sensitive secondary competitive transplantation assays with the same number of WT and *Ret*<sup>–/–</sup> bone marrow cells (Fig. 2a). We found minute frequencies of *Ret*<sup>–/–</sup> cells in the blood (Fig. 2d and Extended Data Fig. 2h), a defect already established in bone marrow LSKs (Fig. 2e). This major impact in long-term transplantation was in contrast to the modest reduction in the potential of short-term CFU-spleen (CFU-s) of *Ret*<sup>–/–</sup> progenitors (Extended Data Fig. 2i). Kidney and enteric nervous system development rely on activation of RET by GDNF and its co-receptor GDNF family receptor alpha 1 (GFRα1) (ref. 7). Analysis of E14.5 *Gfra1*<sup>–/–</sup>, which similarly to *Ret*<sup>–/–</sup> animals die perinatally due to kidney aplasia, revealed normal progenitors and transplantation potential, suggesting that the *Ret*<sup>–/–</sup> HSC phenotype is not secondary to kidney or nervous system deficits (Extended Data Fig. 3). To address HSC-autonomous effects further, we generated *Ret*<sup>fl</sup> mice that were bred to *Vav1*-iCre mice (Extended Data Fig. 4a)<sup>8,15</sup>. Deletion of the *Ret*<sup>fl</sup> allele was inefficient in E14.5 HSCs<sup>16</sup>, and *Vav1*-iCre.*Ret*<sup>fl</sup> mice had normal fetal haematopoiesis (Extended Data Fig. 4b, c). In contrast, effective *Ret* conditional ablation led to reduced adult HSCs (Fig. 2f). In agreement, *Vav1*-iCre.*Ret*<sup>fl</sup> mice died promptly upon 5-fluorouracil treatment (Fig. 2g). Primary and secondary transplantations also revealed that *Vav1*-iCre.*Ret*<sup>fl</sup> progenitors lost their multi-lineage transplantation fitness (Fig. 2h and Extended Data Fig. 4d, e). Altogether, these results indicate that RET is a cell-autonomous requirement to HSC maintenance

<sup>1</sup>Instituto de Medicina Molecular, Faculdade de Medicina de Lisboa, Avenida Professor Egas Moniz, Edifício Egas Moniz, 1649-028 Lisboa, Portugal. <sup>2</sup>Division of Molecular Immunology, MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK. <sup>3</sup>Institute of Immunology and Infection Research, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK.

<sup>4</sup>Laboratory for Neuronal Differentiation and Regeneration, RIKEN Center for Developmental Biology, Kobe 650-0047, Japan. <sup>5</sup>Graduate School of Medicine, Kobe University 7-5-1 Kusunoki-cho, Chuo-ku, Kobe City, Hyogo 650-0017, Japan.

\*These authors contributed equally to this work.

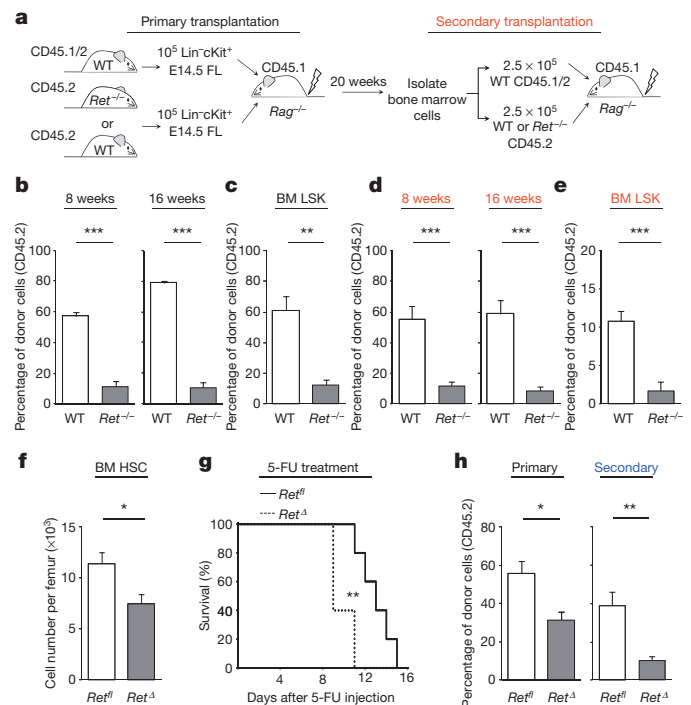


**Figure 1 | *Ret* deficiency leads to reduced HSCs with impaired transplantation potential.** **a**, E14.5 fetal liver HSCs and multipotent progenitor cells, double-positive thymocytes (DP), lymphoid tissue initiator cells (LTin) and adult enteric tissue (ET) were analysed by quantitative PCR with reverse transcription (RT-qPCR). \*\*\**P* value for one-way analysis of variance (ANOVA) lower than 0.001. **b**, Flow cytometry of E14.5 fetal liver HSCs, LTin cells and double-positive thymocytes. Grey, isotype control. **c**, E14.5 fetal liver TER119<sup>+</sup> CD45<sup>+</sup> CD31<sup>+</sup> endothelial cells (EC), TER119<sup>+</sup> CD45<sup>+</sup> CD31<sup>+</sup> cKit<sup>+</sup> ICAM-1<sup>+</sup> hepatocyte progenitors (HP) and TER119<sup>+</sup> CD45<sup>+</sup> CD31<sup>+</sup> cKit<sup>+</sup> ICAM-1<sup>+</sup> mesenchymal cells (MC) were FACS-sorted and analysed by confocal microscopy. **d**, **e**, Confocal microscopy analysis. Arrows, Lin<sup>+</sup> CD150<sup>+</sup> CD48<sup>+</sup> CD41<sup>+</sup> candidate HSCs. Scale bar, 5 μm. **f**, Fetal liver E14.5 LSKs. WT *n* = 20; *Ret*<sup>-/-</sup> *n* = 18. **g**, CFU per 5 × 10<sup>3</sup> Lin<sup>+</sup> cKit<sup>+</sup> cells on day 8. BFU-E, burst-forming unit-erythrocyte; CFU-GM, colony-forming unit-granulocyte/macrophage; CFU-GEMM, colony-forming unit-granulocyte/erythrocyte/macrophage/megakaryocyte. WT *n* = 3; *Ret*<sup>-/-</sup> *n* = 3. **h**, Annexin V<sup>+</sup> cells in cultured LSKs. WT *n* = 7; *Ret*<sup>-/-</sup> *n* = 4. **i**, E14.5 HSCs. WT *n* = 20; *Ret*<sup>-/-</sup> *n* = 18. **j**, HSC cell cycle. WT *n* = 7; *Ret*<sup>-/-</sup> *n* = 6. **k**, Survival upon transplantation. \**P* value for log rank test lower than 0.05. WT *n* = 5; *Ret*<sup>-/-</sup> *n* = 6. Error bars, s.e.m. \* and \*\*, *P* values for Student's *t*-test lower than 0.05 and 0.01 respectively.

and haematopoietic stress responses, a finding also supported by *Ret* upregulation after irradiation-induced genotoxic stress (Extended Data Fig. 5a).

Previous reports have identified a gene signature associated with long-term HSC activity<sup>17–19</sup>. While most of those genes were not significantly modified, *Bcl2* and *Bcl2l1* were heavily reduced in *Ret* null LSKs and HSCs (Fig. 3a, b and Extended Data Fig. 5b, c). The marked reduction of *BCL2* and *BCLxL*, together with the susceptibility of *Ret*<sup>-/-</sup> progenitors to apoptosis, suggested that GFLs could provide survival signals to HSCs. In agreement, GFLs increased blood progenitor survival and preserved HSCs in culture conditions (Fig. 3c, d and Extended Data Fig. 5d).

RET activation in neurons leads to ERK1/2, PI3K/Akt and p38/MAP kinase activation<sup>7</sup>, while phosphorylation of CREB induces *Bcl2* gene

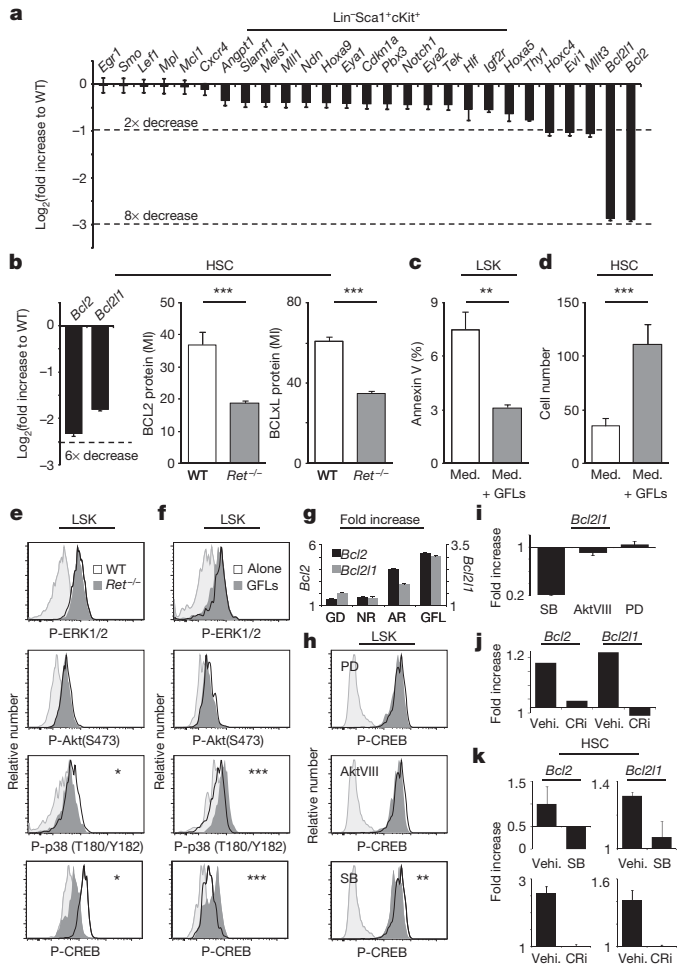


**Figure 2 | RET cell-autonomous signals control haematopoietic stress responses.** **a**, Primary and secondary (red) transplantation. FL, fetal liver. **b**, Donor CD45.2 blood cells in primary transplantations. WT *n* = 5; *Ret*<sup>-/-</sup> *n* = 9. Representative of two independent experiments. **c**, Donor bone marrow (BM) CD45.2 LSK cells. WT *n* = 3; *Ret*<sup>-/-</sup> *n* = 4. **d**, Donor CD45.2 cells after serial transplantation. WT *n* = 5; *Ret*<sup>-/-</sup> *n* = 4. Representative of two independent experiments. **e**, Donor bone marrow CD45.2 LSK after serial transplantation. WT *n* = 5; *Ret*<sup>-/-</sup> *n* = 4. **f**, Bone marrow HSCs in *Ret*<sup>fl</sup> and *Ret*<sup>fl</sup> littermate controls. *Ret*<sup>fl</sup> *n* = 10; *Ret*<sup>fl</sup> *n* = 7. **g**, Survival after treatment with 5-fluorouracil (5-FU). *Ret*<sup>fl</sup> *n* = 5; *Ret*<sup>fl</sup> *n* = 5. \*\**P* value for log rank test lower than 0.01. **h**, Donor blood cells 16 weeks after primary and secondary (blue) transplantation. Primary: *Ret*<sup>fl</sup> *n* = 5; *Ret*<sup>fl</sup> *n* = 4; secondary: *Ret*<sup>fl</sup> *n* = 7; *Ret*<sup>fl</sup> *n* = 9. Similar results were obtained in two independent experiments. Error bars, s.e.m. \*, \*\* and \*\*\*, *P* values for Student's *t*-test lower than 0.05, 0.01 and 0.001 respectively.

family expression<sup>20,21</sup>. Analysis of p38/MAP kinase and CREB in *Ret*<sup>-/-</sup> LSKs revealed that these molecules were hypo-phosphorylated, while ERK1/2 and PI3K/Akt activation was unperturbed (Fig. 3e and Extended Data Fig. 5e). Accordingly, RET activation by GFLs led to rapid p38/MAP kinase and CREB phosphorylation and increased *Bcl2/Bcl2l1* expression by LSKs, while ERK1/2, PI3K/Akt phosphorylation was stable (Fig. 3f, g and Extended Data Fig. 5f). Importantly, inhibition of p38/MAP kinase upon GFL activation led to impaired CREB phosphorylation and *Bcl2/Bcl2l1* expression while inhibition of ERK1/2 and PI3K/Akt had no impact on these molecules (Fig. 3h, i and Extended Data Fig. 5g). Finally, inhibition of CREB upon GFL activation resulted in decreased *Bcl2/Bcl2l1* levels (Fig. 3j). Altogether, these data demonstrate that RET-deficient LSKs express reduced *Bcl2* and *Bcl2l1*, downstream of impaired p38/MAP kinase and CREB activation; this is a signalling pathway that was also operational in purified HSCs (Fig. 3k).

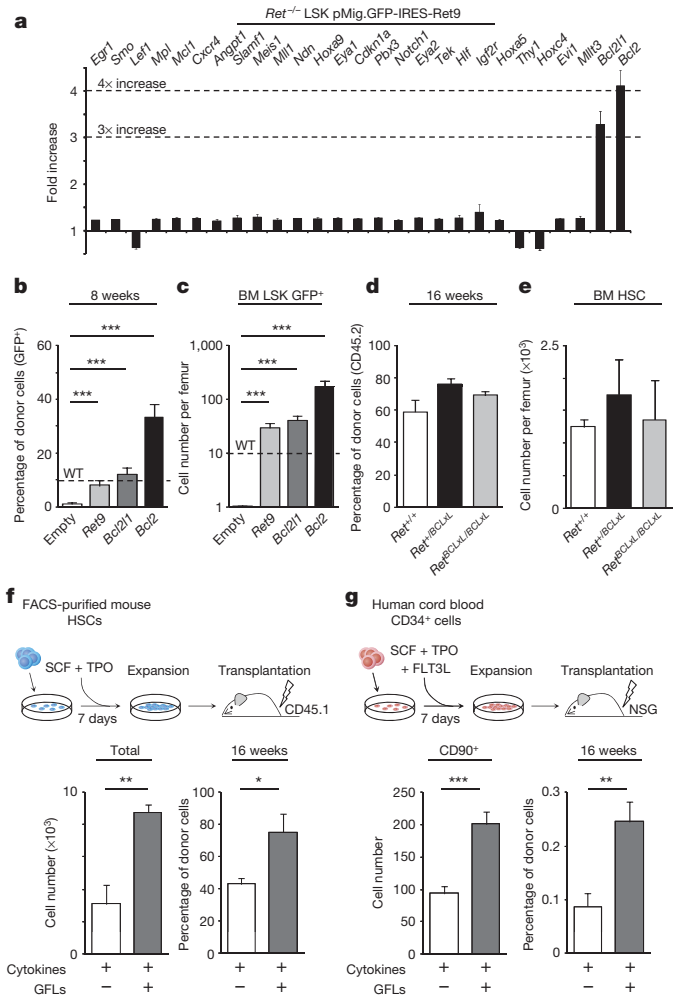
These findings suggested that reduced *Bcl2* and *Bcl2l1* caused the unfitness of *Ret*-deficient HSC. Retroviral transductions showed that *Bcl2* and *Bcl2l1* expression levels were quickly restored in *Ret*<sup>-/-</sup> LSKs transduced with WT *Ret*, while other signature genes were unperturbed (Fig. 4a). To test whether *Ret*<sup>-/-</sup> progenitor fitness could be restored by enforced *Ret* expression, we performed competitive transplantations with *Ret*<sup>-/-</sup> progenitors transduced with pMig.*Ret*9.IRES-GFP (IRES, internal ribosomal entry site; GFP, green fluorescent protein) retrovirus together with competitor CD45.1 bone marrow (Extended Data Fig. 6a). Restoration of RET expression fully rescued *Ret*<sup>-/-</sup> progenitor transplantation (Fig. 4b, c), and enforced expression of RET downstream





**Figure 3 | RET induces *Bcl2/Bcl2l1* downstream of p38/MAP and CREB activation.** **a**, Fetal liver E14.5 *Ret*<sup>-/-</sup> and WT LSKs (*n* = 3). **b**, Left, fetal liver E14.5 *Ret*<sup>-/-</sup> and WT HSCs (*n* = 3). Right, mean intensity (MI) for BCL2 and BCLxL protein in fetal liver E14.5 *Ret*<sup>-/-</sup> and WT HSCs. BCL2: WT *n* = 40, *Ret*<sup>-/-</sup> *n* = 29; BCLxL: WT *n* = 40, *Ret*<sup>-/-</sup> *n* = 40. **c**, Annexin V<sup>+</sup> LSK cells cultured in medium (Med.) with or without GFLs (*n* = 9). **d**, HSC numbers after GFL treatment (*n* = 12). **e**, E14.5 *Ret*<sup>-/-</sup> and WT littermate control LSKs. P-ERK and P-Akt: WT *n* = 15, *Ret*<sup>-/-</sup> *n* = 16; P-p38: WT *n* = 3, *Ret*<sup>-/-</sup> *n* = 3; P-CREB: WT *n* = 16, *Ret*<sup>-/-</sup> *n* = 19. **f**, LSK activation by GFLs. P-ERK and P-Akt: alone *n* = 12, GFLs *n* = 12; P-p38: alone *n* = 12, GFLs *n* = 12; P-CREB: alone *n* = 6, GFLs *n* = 6. **g**, LSKs upon treatment with GDNF (GD), NRTN (NR), ARTN (AR) and three GFLs (GFL). Relative to PBS-treated LSKs (vehicle). **h**, LSKs cultured with GFLs (black line) or GFLs and the inhibitors SB 202190 (SB), PD98,059 (PD) or Akt1/2, Akt Inhibitor VIII (AktVIII) (solid grey). PD, AktVIII and SB: GFLs *n* = 6, GFLs+inh *n* = 6. **i**, LSK cells upon GFL treatment. Relative to GFL-treated LSKs. **j**, Treatment with GFLs + DMSO (vehicle (Vehi.)) or GFLs + CBP-CREB interaction inhibitor (CRi), relative to DMSO-treated LSKs (vehicle). **k**, HSCs after treatment with GFL + DMSO (vehicle), GFL + SB 202190 (SB) or GFL + CBP-CREB interaction inhibitor (CRi), relative to DMSO-treated HSCs (vehicle). Error bars, s.e.m. Light grey, isotype control. \*\* and \*\*\*, *P* values for Student's *t*-test lower than 0.01 and 0.001 respectively.

targets, *Bcl2* or *Bcl2l1*, was sufficient to recover the engraftment of *Ret* null LSKs but had no effect on their WT counterparts (Fig. 4b, c and Extended Data Fig. 6b). Since retroviral expression became unstable *in vivo*, we generated a knock-in mouse expressing *BCL2L1* under the control of *Ret* (*Ret*<sup>BCLxL</sup>) (Extended Data Fig. 7a). *Ret*<sup>BCLxL</sup> heterozygous mice had increased *BCL2L1* expression but displayed normal haematopoiesis and HSCs (Extended Data Fig. 7b, c). *Ret*-deficient mice expressing ectopic *BCL2L1* (*Ret*<sup>BCLxL/BCLxL</sup>) had normal fetal progenitors (Extended Data Fig. 7b, d) and transplantation of these progenitors was similar to WT levels at week 16 (Fig. 4d, e and Extended Data Fig. 7e). These data



**Figure 4 | RET activation promotes HSC expansion and transplantation efficiency.** **a**, Fold increase between *Ret*<sup>-/-</sup> pMig-GFP-IRES-*Ret9* and empty vector (*n* = 3). **b**, CD45.2<sup>+</sup>GFP<sup>+</sup> blood cells 8 weeks after transplantation with *Ret*<sup>-/-</sup> progenitors transduced with pMig-GFP-IRES-Empty (Empty) virus, or expressing *Ret9*, *Bcl2l1* or *Bcl2*. Empty *n* = 13; *Ret9* *n* = 5; *Bcl2l1* *n* = 8; *Bcl2* (*n* = 7). **c**, Bone marrow LSKs. *Ret*<sup>-/-</sup> Empty *n* = 13; *Ret9* *n* = 4; *Bcl2l1* *n* = 6; *Bcl2* (*n* = 5). **d**, Donor CD45.2 blood cells 16 weeks after competitive transplantation. *Ret*<sup>+/-</sup> (*n* = 8); *Ret*<sup>+/-</sup>/*BCLxL* (*n* = 8); *Ret*<sup>BCLxL/BCLxL</sup> (*n* = 8). Similar results were obtained in two independent experiments. **e**, Bone marrow CD45.2 HSCs 16 weeks after transplantation. *Ret*<sup>+/-</sup> *n* = 3; *Ret*<sup>+/-</sup>/*BCLxL* *n* = 3; *Ret*<sup>BCLxL/BCLxL</sup> *n* = 3. Similar results were obtained in two independent experiments. **f**, Purified HSCs were expanded for 7 days with or without GFLs and were transplanted with competitor CD45.1 bone marrow cells. Bottom left: nucleated cells at culture day 7 (*n* = 11). Bottom right: culture-derived blood cells (CD45.2) at 16 weeks after transplantation (*n* = 5). Similar results were obtained in two independent experiments. **g**, Top: scheme expansion and transplantation of human CD34<sup>+</sup> cells. Bottom left: CD34<sup>+</sup>CD38<sup>-</sup>CD90<sup>+</sup> cells at culture day 7 (*n* = 20). Bottom right: culture-derived hCD45<sup>+</sup> blood cells at 16 weeks after transplantation (*n* = 5). Similar results were obtained in two independent experiments. Error bars, s.e.m. \*, \*\*, and \*\*\*, *P* values for Student's *t*-test lower than 0.05, 0.01 and 0.001 respectively.

indicate that *Bcl2l1* has no significant impact in WT progenitors but is sufficient to restore the function of *Ret* null progenitors.

Evidence that RET signals promote HSC expansion was provided by *in vitro* cultures of fluorescence-activated cell sorting (FACS)-purified HSCs with stem cell factor, thrombopoietin and GFLs. Addition of GFLs to purified HSCs improved their expansion (Fig. 4f); and transplantation of expanded HSCs revealed that GFL-treated HSCs had increased fitness (Fig. 4f). Remarkably, addition of GFLs to human cord blood CD34<sup>+</sup> progenitors significantly increased *BCL2* and *BCL2L1* expression; improved expansion of primitive progenitors and resulted in increased

repopulation activity of human progenitors into NOD.Cg-*Prkdc<sup>scid</sup> Il2rg<sup>tm1Wjl</sup>/SzJ* (NSG) mice (Fig. 4g and Extended Data Fig. 8).

Our results reveal that RET signalling is a crucial cell-autonomous pathway controlling fetal and adult HSC survival via BCL2 family members. Previous studies have identified molecules that co-regulate HSC maintenance and differentiation<sup>1,2,17–19</sup>. We now show that RET activation by GFLs specifically regulates HSC survival, preserving HSC stemness and discriminating between HSC maintenance and progenitor differentiation. Despite *Bcl2* expression by HSCs, no appreciable LSK transplantation deficiencies were reported in *Bcl2*-deficient mice, although reduced haematopoietic precursors were reported in *Bcl2l1<sup>-/-</sup>* mice<sup>22,23</sup>. Thus, *Bcl2* and *Bcl2l1* may have redundant roles in HSCs, an idea supported by our data demonstrating that *Bcl2l1* or *Bcl2* are independently sufficient to rescue *Ret<sup>-/-</sup>* HSCs (Fig. 4b–e). Contrary to nervous cells, we suggest that HSCs use GFLs in a redundant manner<sup>11</sup> since analysis of RET co-receptor single knockouts revealed normal LSKs (Extended Data Fig. 3 and Extended Data Fig. 9)<sup>24–26</sup>.

Our study indicates that absence of neurotrophic factor cues leads to impaired HSC survival and transplantation. Accordingly, activation of RET results in improved HSC survival, expansion and *in vivo* transplantation efficiency. Thus, we propose that RET controls HSC response to physiological demands (Extended Data Fig. 10). Altogether, these findings open new horizons for pre-clinical testing of GFLs in human haematopoietic progenitor expansion and transplantation.

Previous work revealed that nervous cells modulate HSC function<sup>3–5,27–29</sup>; we now show that HSCs are direct targets for neurotrophic factors, indicating that HSCs and neurons are regulated by similar signals. Thus, our work puts forward a possible regulation of neuronal activity by primitive blood progenitors through neurotrophic factor consumption in the HSC environment.

## METHODS SUMMARY

Mice were bred and maintained at the Instituto de Medicina Molecular animal facility. Lin<sup>-</sup>cKit<sup>+</sup> cells were MACS (Miltenyi Biotec) sorted and injected alone or in direct competition with a third-party WT competitor CD45.1/CD45.2 into lethally irradiated CD45.1 mice. Secondary reconstitution experiments were performed on FACS-sorted bone marrow cells from primary recipients and injected intravenously with third-party cells. FACS-sorted murine HSCs were cultured for 7 days in StemSpan SFEM (STEMCELL Technologies). Human cord blood CD34<sup>+</sup> cells were cultured similarly to murine HSCs, with added rmFLT3.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

**Received 27 June 2012; accepted 21 May 2014.**

**Published online 27 July 2014.**

1. Trumpp, A., Essers, M. & Wilson, A. Awakening dormant haematopoietic stem cells. *Nature Rev. Immunol.* **10**, 201–209 (2010).
2. Morrison, S. J. & Scadden, D. T. The bone marrow niche for haematopoietic stem cells. *Nature* **505**, 327–334 (2014).
3. Yamazaki, S. *et al.* Nonmyelinating Schwann cells maintain hematopoietic stem cell hibernation in the bone marrow niche. *Cell* **147**, 1146–1158 (2011).
4. Mendez-Ferrer, S., Lucas, D., Battista, M. & Frenette, P. S. Haematopoietic stem cell release is regulated by circadian oscillations. *Nature* **452**, 442–447 (2008).
5. Lucas, D. *et al.* Chemotherapy-induced bone marrow nerve injury impairs hematopoietic regeneration. *Nature Med.* **19**, 695–703 (2013).
6. Veiga-Fernandes, H. *et al.* Tyrosine kinase receptor RET is a key regulator of Peyer's Patch organogenesis. *Nature* **446**, 547–551 (2007).
7. Mulligan, L. M. RET revisited: expanding the oncogenic portfolio. *Nature Rev. Cancer* **14**, 173–186 (2014).
8. Almeida, A. R. *et al.* RET/GFR $\alpha$  signals are dispensable for thymic T cell development *in vivo*. *PLoS ONE* **7**, e52949 (2012).

9. Kiel, M. J., Yilmaz, O. H., Iwashita, T., Terhorst, C. & Morrison, S. J. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* **121**, 1109–1121 (2005).
10. Kim, I., He, S., Yilmaz, O. H., Kiel, M. J. & Morrison, S. J. Enhanced purification of fetal liver hematopoietic stem cells using SLAM family receptors. *Blood* **108**, 737–744 (2006).
11. Patel, A. *et al.* Differential RET signaling pathways drive development of the enteric lymphoid and nervous systems. *Sci. Signal.* **5**, ra55 (2012).
12. Winkler, I. G. *et al.* Positioning of bone marrow hematopoietic and stromal cells relative to blood flow *in vivo*: serially reconstituting hematopoietic stem cells reside in distinct nonperfused niches. *Blood* **116**, 375–385 (2010).
13. Ding, L. & Morrison, S. J. Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature* **495**, 231–235 (2013).
14. Schuchardt, A., D'Agati, V., Larsson-Blomberg, L., Costantini, F. & Pachnis, V. Defects in the kidney and enteric nervous system of mice lacking the tyrosine kinase receptor Ret. *Nature* **367**, 380–383 (1994).
15. de Boer, J. *et al.* Transgenic mice with hematopoietic and lymphoid specific expression of Cre. *Eur. J. Immunol.* **33**, 314–325 (2003).
16. McMahon, K. A. *et al.* Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. *Cell Stem Cell* **1**, 338–345 (2007).
17. Mansson, R. *et al.* Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. *Immunity* **26**, 407–419 (2007).
18. Terskikh, A. V., Miyamoto, T., Chang, C., Diatchenko, L. & Weissman, I. L. Gene expression analysis of purified hematopoietic stem cells and committed progenitors. *Blood* **102**, 94–101 (2003).
19. Thompson, B. J. *et al.* Control of hematopoietic stem cell quiescence by the E3 ubiquitin ligase Fbw7. *J. Exp. Med.* **205**, 1395–1408 (2008).
20. Perianayagam, M. C., Madias, N. E., Pereira, B. J. & Jaber, B. L. CREB transcription factor modulates Bcl2 transcription in response to C5a in HL-60-derived neutrophils. *Eur. J. Clin. Invest.* **36**, 353–361 (2006).
21. Shukla, A. *et al.* Activated cAMP response element binding protein is overexpressed in human mesotheliomas and inhibits apoptosis. *Am. J. Pathol.* **175**, 2197–2206 (2009).
22. Nakayama, K., Negishi, I., Kuida, K., Sawa, H. & Loh, D. Y. Targeted disruption of Bcl-2 $\alpha\beta$  in mice: occurrence of gray hair, polycystic kidney disease, and lymphocytopenia. *Proc. Natl Acad. Sci. USA* **91**, 3700–3704 (1994).
23. Motoyama, N. *et al.* Massive cell death of immature hematopoietic cells and neurons in Bcl-x-deficient mice. *Science* **267**, 1506–1510 (1995).
24. Cacalano, G. *et al.* GFR $\alpha$ 1 is an essential receptor component for GDNF in the developing nervous system and kidney. *Neuron* **21**, 53–62 (1998).
25. Nishino, J. *et al.* GFR  $\alpha$ 3, a component of the artemin receptor, is required for migration and survival of the superior cervical ganglion. *Neuron* **23**, 725–736 (1999).
26. Rossi, J. *et al.* Retarded growth and deficits in the enteric and parasympathetic nervous system in mice lacking GFR  $\alpha$ 2, a functional neurturin receptor. *Neuron* **22**, 243–252 (1999).
27. Katayama, Y. *et al.* Signals from the sympathetic nervous system regulate hematopoietic stem cell egress from bone marrow. *Cell* **124**, 407–421 (2006).
28. Mendez-Ferrer, S. *et al.* Mesenchymal and haematopoietic stem cells form a unique bone marrow niche. *Nature* **466**, 829–834 (2010).
29. Spiegel, A. *et al.* Catecholaminergic neurotransmitters regulate migration and repopulation of immature human CD34<sup>+</sup> cells through Wnt signaling. *Nature Immunol.* **8**, 1123–1131 (2007).

**Acknowledgements** We thank I. Monteiro Grillo and the radiotherapy service at Hospital de Santa Maria; H. Ferreira and the service of obstetrics, gynaecology and reproductive medicine at the Hospital of Santa Maria; the Instituto de Medicina Molecular animal facility, flow cytometry unit, bioimaging unit and histology unit for technical assistance. We also thank all members of H.V.-F. laboratory for discussion. D.F.-P., S.A.-M., R.G.D. and A.R.M.A. were supported by scholarships from Fundação para a Ciência e Tecnologia, Portugal. H.V.-F. was supported by Fundação para a Ciência e Tecnologia (PTDC/SAU-MII/104931/2008), Portugal, the European Molecular Biology Organisation (Project 1648), European Research Council (Project 207057) and National Blood Foundation, USA.

**Author Contributions** D.F.-P., S.A.-M., M.R.-C., I.B., R.G.D., T.B., A.R.M.A. and H.R. did experiments and data analysis; H.E. generated *Ret<sup>ECXL</sup>* mice; D.F.-P., S.A.-M., A.P. and H.V.-F. designed *in vivo* and *ex vivo* experiments; D.F.-P. and H.V.-F. wrote the manuscript and H.V.-F. directed the study.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.V.-F. ([jhfernandes@medicina.ulisboa.pt](mailto:jhfernandes@medicina.ulisboa.pt)).

## METHODS

**Mice.** C57BL/6J (CD45.2 and CD45.1), *Rag1*<sup>-/-</sup> (CD45.2 and CD45.1)<sup>30</sup>, *Vav1*Cre<sup>15</sup>, *Gfra1*<sup>-/-</sup> (ref. 24), *Gfra2*<sup>-/-</sup> (ref. 26), *Gfra3*<sup>-/-</sup> (ref. 25), *Ret*<sup>-/-</sup> (ref. 14) and *Ret*<sup>fl</sup> (ref. 8) were maintained at the Instituto de Medicina Molecular. NSG mice were bought from The Jackson Laboratory. *Ret*<sup>BCLxL-IRES-Puro</sup> (*Ret*<sup>BCLxL</sup>) knock-in mice were generated by inserting a gene cassette composed of human *BCL2L1* complementary DNA (cDNA) followed by ires-puromycin resistance gene into the endogenous *Ret* locus via gene targeting as describe previously<sup>31</sup>. We performed power analysis to determine sample size. No exclusion, blinding or randomization criteria were used in experiments involving genetically modified animals. Mice were systematically compared with sex-matched littermate controls. All mice strains were bred and maintained at the Instituto de Medicina Molecular animal facility. Animal procedures were in accordance with national and institutional guidelines.

**Colony-forming unit assays and homing capacity.** Five thousand E14.5 Lin<sup>-</sup>cKit<sup>+</sup> cells were MACS purified (Miltenyi Biotec) from WT and *Ret*<sup>-/-</sup> and cultured in M3434 (Stem Cell Technologies), and scored at days 8–10 by flow cytometry and microscope analysis. Thirty thousand E14.5 Lin<sup>-</sup>cKit<sup>+</sup> cells were MACS (Miltenyi Biotec) purified from WT or *Ret*<sup>-/-</sup>, injected into lethally irradiated mice (9 Gy) and CFU-s scored at day 12 by microscope analysis. Homing assays used E14.5 Lin<sup>-</sup>cKit<sup>+</sup> cells labelled with CMTMR (Invitrogen), injected into lethally irradiated mice. We performed flow cytometry analysis 20 h after injection.

**Transplantation experiments.** For reconstitution experiments with fetal liver,  $1 \times 10^5$  E14.5 Lin<sup>-</sup>cKit<sup>+</sup> cells were MACS sorted from WT, *Ret*<sup>-/-</sup>, *Ret*<sup>+/BCLxL</sup> or *Ret*<sup>BCLxL/BCLxL</sup> and were injected alone or in direct competition with a third-party WT competitor CD45.1/CD45.2 (1:1 ratio) into lethally irradiated *Rag1*<sup>-/-</sup> CD45.1 mice. For secondary reconstitution experiments with bone marrow,  $2.5 \times 10^5$  cells from each genotype were FACS-sorted from primary recipients and injected intravenously in competition with the WT CD45.1/CD45.2 third-party competitor cells into lethally irradiated *Rag1*<sup>-/-</sup> CD45.1 mice. *Vav1*-iCre.*Ret*<sup>fl</sup> competitive transplants were analysed at 16 weeks after transplantation of  $2.5 \times 10^5$  cells from each genotype together with CD45.1/CD45.2 competitor WT cells.

**Rescue of *in vivo* transplantation.** E14.5 Lin<sup>-</sup>cKit<sup>+</sup>Sca1<sup>+</sup> WT or *Ret*<sup>-/-</sup> cells were transduced overnight with pMig.IRES-GFP retroviral vector containing *Ret9*, *Bcl2* or *Bcl2l1*. GFP<sup>+</sup> cells were further FACS purified for immediate transcriptional analysis. Lin<sup>-</sup>cKit<sup>+</sup>GFP<sup>+</sup> cells were injected into lethally irradiated mice. Six weeks later transduced bone marrow Lin<sup>-</sup>CD45.2<sup>+</sup>GFP<sup>+</sup> cells were purified by flow cytometry and  $10^5$  cells were co-injected with a radio-protective dose of  $10^5$  CD45.1 bone marrow cells into lethally irradiated recipients.

**5-Fluorouracil treatment.** *Vav1*-iCre.*Ret*<sup>fl</sup> and their littermate controls were injected weekly with 150 µg of 5-fluorouracil per gram of body weight.

**Flow cytometry.** Embryonic fetal livers were micro-dissected and homogenized in 70 µm cell strainers. Bone marrow cells were either collected by flushing or crushing bones. Bone marrow cells numbers were calculated per femur. Cell suspensions were stained with: anti-CD117 (cKit) (2B8), anti-Ly-6A/E (Sca-1) (D7), anti-CD16/32 (FcγRII/III) (93), anti-CD3 (eBio500A2), anti-CD150 (mShad150), anti-CD48 (HM48-1), anti-CD19 (eBio1D3), anti-CD11b (M1/70), anti-Ly-6G (Gr-1) (RB6-8C5), anti-Ly79 (TER119), anti-NK1.1 (PK136), anti-CD11c (N418), anti-CD45.1 (A20), anti-CD45.2 (104), anti-CD54 (ICAM-1) (YNI1.7.4), anti-CD34 (RAM34), anti-CD51 (RMV-7) and anti-CD41 (eBioMWReg30) from eBioscience; anti-CD38 (90), anti-CD3 (145-2C11), anti-CD34 (HM34) and anti-CD31 (390) from BioLegend; anti-Ly6C (HK1.4) from Abcam, Annexin V from BD Pharmingen. Lineage cocktail include anti-CD3, anti-CD19, anti-Ly-6G, anti-Ly6C, anti-Ly79, anti-NK1.1, anti-CD11c for embryonic fetal livers plus anti-CD11b for adult bone marrow cells. HSCs were defined as Lin<sup>-</sup>cKit<sup>+</sup>CD150<sup>+</sup>CD48<sup>-</sup> cells, LSKs as Lin<sup>-</sup>Sca1<sup>+</sup>cKit<sup>+</sup> cells and myeloid progenitors were defined as Lin<sup>-</sup>Sca1<sup>-</sup>cKit<sup>+</sup> cells. Human cord blood was enriched for CD34<sup>+</sup> cells using CD34 MicroBead Kit (Miltenyi Biotec) after Histopaque separation (Sigma) and stained with anti-human CD34 (AC136) (Miltenyi Biotec), anti-human CD38 (HIT2) (eBioscience), anti-human CD45 (Hi30) (eBioscience) and anti-human CD90 (5E10) (eBioscience). Samples were sorted on a FACSAria I or FACSAria III and analysed on a FACSCanto or LSRFortessa (BD). We analysed flow cytometry data with FlowJo 8.8.7 software (Tree Star).

**Immunofluorescence of sorted cells.** Two thousand to forty thousand cells were seeded in poly-lysine coated coverslips (Sigma Aldrich P8920). Cells were fixed for 15 min in 2% PFA at 4 °C or 10 min with methanol at -20 °C (for BCLxL staining). Slides were then washed with PBS, permeabilized with 0.1% Triton X-100 for 10 min at 4 °C and stained in 3% FBS for 45 min at 4 °C with anti-RET (Neuromics GT15002), anti-GDNF (Abcam 18956), anti-Neurturin (R&D Systems AF477), anti-Artemin (R&D Systems 185234), anti-mouse BCL2 (3F11) (BD Pharmingen) or anti-BCLxL (H-62) (Santa Cruz biotechnology). Secondary antibody staining plus DAPI was done for 30 min at 4 °C using anti-rabbit (Invitrogen A21206), anti-mouse (Invitrogen A21127), anti-goat (Invitrogen A11078), anti-rat (Invitrogen A11006) or anti-Armenian hamster (Jackson ImmunoResearch 127-165-099). Slides were

mounted in Mowiol (Calbiochem), images acquired on a Zeiss LSM 710 (×63 objective) and images analysed using ImageJ software.

**Histology and immunofluorescence.** Whole-mount bone marrow samples were prepared as previously described<sup>32</sup>. Briefly, sternal bones were collected and transected with a surgical blade into two or three fragments. Fragments were bisected sagittally for the bone marrow cavity to be exposed, fixed in 4% PFA and blocked and permeabilized in 1× PBS with 2% BSA, 10% FBS, 0.6% Triton X-100, followed by an Avidin/Biotin Blocking Kit (Vector laboratories).

Frozen section preparation. Femurs were placed in 4%PFA for 2 h at 4 °C and imbued overnight in 30% sucrose. Bones were then included in Cutting Temperature (OCT) compound (Sakura), snap frozen in *N*-methylbutane chilled in liquid nitrogen and kept at -80 °C. Sections (7 µm) obtained using a Cryostat LEICA CM 3050S with a tungsten carbide blade were placed in coated slides. E14.5 fetal liver or E15.5 enteric tissue were placed in 4% PFA overnight at 4 °C followed imbedding in 10% (2 h), 20% (2 h) and 30% sucrose (overnight). Fetal liver or enteric tissue were then included in OCT, frozen in dry ice and kept at -30 °C. Sections (10 µm) obtained using a Cryostat LEICA CM 3050S were placed on coated slides. Slides were air dried, rinsed with PBS and blocked for 30 min at room temperature with 1× PBS, 2% BSA, 10% FBS. Then sections were washed and blocked using an Avidin/Biotin Blocking Kit (Vector laboratories), and permeabilized with PBS 0.3% Triton X-100 for 10 min at room temperature (about 22 °C).

Immunofluorescence staining of whole-mount tissues and frozen sections. Slides or samples were incubated overnight (or for 1–2 days for the whole-mount samples), at 4 °C with biotin-labelled antibodies in PBS: CD3 (eBio500A2), anti-CD19 (eBio1D3), anti-Ly79 (TER119), anti-Ly-6G (Gr-1), anti-CD11c (N418), anti-CD41 (eBioMWReg30) and anti-CD48 (HM48-1); together with anti-CD150 (mShad150) Alexa Fluor 488 and with either one of the following primary antibodies: anti-GDNF (Abcam 18956), anti-Neurturin (R&D Systems AF477) or anti-Artemin (R&D Systems 185234). Samples were then washed in 1× PBS, 2% BSA, 10% FBS and stained with Streptavidin A546-conjugated (Invitrogen S11225) together with either A647 anti-rabbit (Invitrogen A21244), anti-goat (Invitrogen A21447), anti-rat (Invitrogen A21247) or A405 anti-rabbit (abcam ab175652) secondary antibodies plus DAPI for 45 min at room temperature (or 2 ½ h plus TOPRO3 (Invitrogen) for whole-mount samples). Samples were washed with 1× PBS, 2% BSA, 10% FBS and bone marrow, fetal liver and enteric tissue sections were mounted using Mowiol (Calbiochem 475904) while whole-mount samples were dehydrated in methanol and optically cleared using benzyl alcohol:benzyl benzoate (BABB) (Sigma)<sup>6,33</sup>. Sections or whole-mount images were acquired with a Zeiss LSM 710 (×40 or ×63 objective lens for whole-mount or frozen sections respectively) and images were processed using Zeiss LSM Image Browser (Carl Zeiss).

**Intracellular staining.** Intracellular staining used BrdU Flow Kit, anti-P-S6 (pS235/pS236) (N7-548) and anti-P-Akt (pT308) (JI-223.371) from BD Pharmingen, anti-PIP<sub>3</sub> (Z-P345) from Echelon Biosciences, and anti-P-CREB (pS133) (87G3), anti-P-p38 (pT180/Y182) (28B10), anti-P-Akt (pS473) (D9E) and anti-P-ERK1/2 (pT202/pY204) (D13.14.4E) from Cell Signaling Technology. Intracellular staining used anti-human RET (132507) from R&D Systems according to the manufacturer's instructions.

**Signalling and cell death.** One million E14.5 WT Lin<sup>-</sup>cKit<sup>+</sup> cells were cultured in DMEM and starved for 2 h. To test CREB phosphorylation upon GFL stimulation Lin<sup>-</sup>cKit<sup>+</sup> cells were stimulated for 1 h with 500 ng ml<sup>-1</sup> each of GFL and co-receptor (rrGFR-α1, rmGFR-α2, rhGFR-α3 and rrGDNF from R&D Systems; rhNRTN and rhARTN from PeproTech). When referring to the use of 'GFLs', we have employed GDNF, NRTN, ARTN and their specific co-receptors in combination. LSK and HSC cells were purified by flow cytometry and stimulated overnight with GFL/GFR-α combinations to determine *Bcl2* and *Bcl2l1* expression levels. For inhibition experiments cells were incubated 2 h before GFL stimulation, to test CREB phosphorylation, or during overnight stimulation with GFLs, to determine *Bcl2* and *Bcl2l1* expression levels, with SB 202190 and PD98,059 from Sigma-Aldrich or Akt1/2, Akt Inhibitor VIII and CBP-CREB Interaction Inhibitor from Calbiochem. To detect Annexin V,  $4 \times 10^4$  E14.5 WT, *Ret*<sup>-/-</sup>, *Ret*<sup>+/BCLxL</sup> or *Ret*<sup>BCLxL/BCLxL</sup> Lin<sup>-</sup>cKit<sup>+</sup> cells per well were cultured overnight in DMEM alone or with GFL/GFR-α. Lin<sup>-</sup>Sca1<sup>+</sup>cKit<sup>+</sup> cells were stimulated with GFL/GFR-α for 120 h and sequentially analysed by flow cytometry.

**Haematopoietic stem cell expansion and transplantation.** FACS-sorted murine HSCs were cultured for 7 days in StemSpan SFEM (STEMCELL Technologies) with recombinant mSCF (PeproTech), mTPO (R&D Systems) and in the presence or absence of 500 ng ml<sup>-1</sup> each of GFL and co-receptor (rrGFR-α1, rmGFR-α2, rhGFR-α3 and rrGDNF from R&D Systems; rhNRTN and rhARTN from PeproTech). Human cord blood CD34<sup>+</sup> cells were cultured similarly to murine HSCs, adding rmFLT3. Expanded cells were then transplanted in competition with CD45.1 bone marrow into lethally irradiated CD45.1 mice (murine) or into NSG mice irradiated with 250 rad (human). All human samples were obtained with informed consent



and protocols were approved by the Centro Hospitalar Lisboa Norte/Faculdade de Medicina de Lisboa Health Ethics Committee.

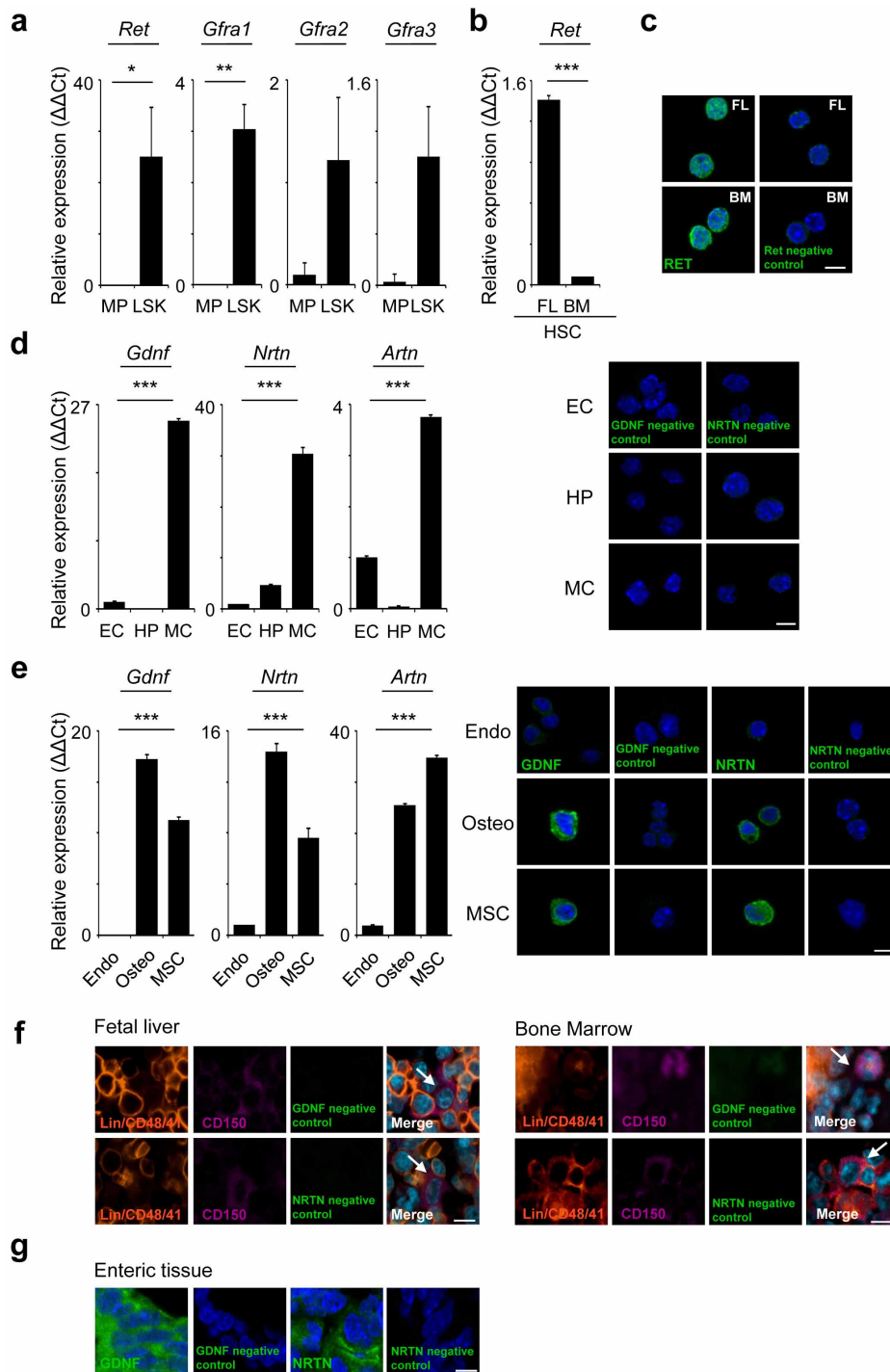
**Real-time PCR analysis.** RNA was extracted from cell suspension using RNeasy Mini Kit or RNeasy Micro Kit (Qiagen). Real-time PCR for *Ret*, *Gfra1*, *Gfra2* and *Gfra3* was done as previously described<sup>6,34</sup>. *Hprt1* was used as housekeeping gene. For TaqMan assays (Applied Biosystems) RNA was retro-transcribed using a High Capacity RNA-to-cDNA Kit (Applied Biosystems), followed by a pre-amplification PCR using TaqMan PreAmp Master Mix (Applied Biosystems). TaqMan Gene Expression Master Mix (Applied Biosystems) was used in real-time PCR. TaqMan Gene Expression Assays bought from Applied Biosystems were the following: *Gapdh* Mm99999915\_g1; *Hprt1* Mm00446968\_m1; *Gusb* Mm00446953\_m1; *Mpl* Mm00440310\_m1; *Mcl1* Mm00725832\_s1; *Meis1* Mm00487664\_m1; *Angpt1* Mm00456503\_m1; *Eya1* Mm00438796\_m1; *Eya2* Mm00802562\_m1; *Egr1* Mm00656724\_m1; *Tek* Mm00443243\_m1; *Slamf1* Mm00443316\_m1; *Lef1* Mm00550265\_m1; *Thy1* Mm00493681\_m1; *Mllt3* Mm00466169\_m1; *Hoxa5* Mm00439362\_m1; *Hoxa9* Mm00439364\_m1; *Hoxc4* Mm00442838\_m1; *Pbx3* Mm00479413\_m1; *Ndn* Mm02524479\_s1; *Evi1* Mm00514814\_m1; *Mlll1* Mm01179213\_g1; *Hlf* Mm00723157\_m1; *Cxcr4* Mm01292123\_m1; *Smo* Mm01162710\_m1; *Igf2r* Mm00439576\_m1; *Cdkn1a* Mm00432448\_m1; *Notch1* Mm00435249\_m1; *Kitl* Mm00442972\_m1; *Thpo* Mm00437040\_m1; *Bcl2l1* Mm00437783\_m1; *Bcl2* Mm00477631\_m1; *Pspn* Mm00436009\_g1; *Artn* Mm00507845\_m1; *Nrtn* Mm03024002\_m1; *Gdnf* Mm00599849\_m1; *Ret* Mm00436304\_m1. For HSC signature gene arrays, gene expression levels were normalized to *Gapdh*, *Hprt1* and *Gusb*. For *Bcl2/Bcl2l1* expression after HSC stimulation

and *Ret* expression levels after *in vivo* transfer, gene expression levels were normalized to *Gapdh* and *Hprt1*. Real-time PCR for human samples used the following primers: *GAPDH* AGGTGAAGGTCGGAGTCAAC and TCTCCATGGTGGTG AAGACG; *BCL2* GCACCTGCACACCTGGAT and CCAAAGTGCAGAGTCTTCAG; *BCL2L1* AGCCTTGGATCCAGGAGAAC and AGCGGTTGAAGCGTTCCT.

**Statistics.** Statistical analysis used Microsoft Excel. Variance was analysed using *F*-test. Student's *t*-test was performed on homocedastic populations, and Student's *t*-test with Welch correction was applied on samples with different variances. When comparing more than two samples, one-way ANOVA was employed. We analysed Kaplan–Meier survival curves with a log rank test. \*, \*\* and \*\*\* represent *P* values lower than 0.05, 0.01 and 0.001, respectively.

30. Mombaerts, P. *et al.* RAG-1-deficient mice have no mature B and T lymphocytes. *Cell* **68**, 869–877 (1992).
31. Uesaka, T. & Enomoto, H. Neural precursor death is central to the pathogenesis of intestinal aganglionosis in *Ret* hypomorphic mice. *J. Neurosci.* **30**, 5211–5218 (2010).
32. Kunisaki, Y. *et al.* Arteriolar niches maintain haematopoietic stem cell quiescence. *Nature* **502**, 637–643 (2014).
33. van de Pavert, S. A. *et al.* Maternal retinoids control type 3 innate lymphoid cells and set the offspring immunity. *Nature* **508**, 123–127 (2014).
34. Peixoto, A. *et al.* CD8 single-cell gene coexpression reveals three different effector types present at distinct phases of the immune response. *J. Exp. Med.* **204**, 1193–1205 (2007).

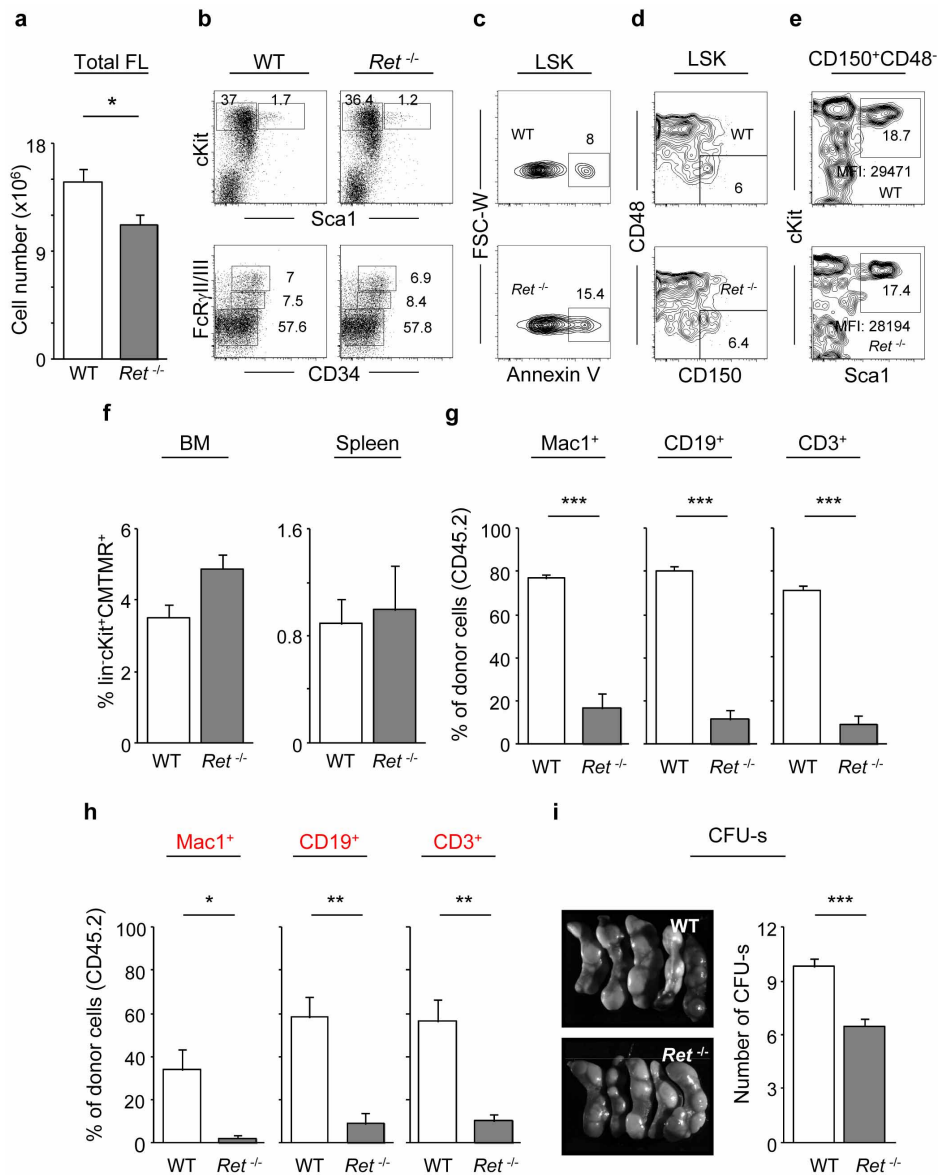




#### Extended Data Figure 1 | *Ret* expression in haematopoietic progenitors and *Ret* ligand expression in the fetal and adult HSC environment.

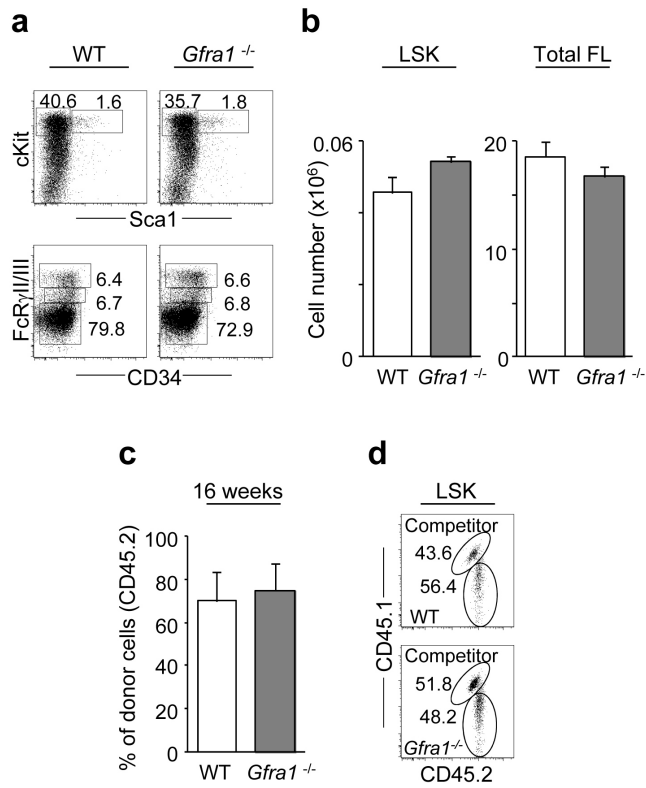
**a**, FACS-sorted E14.5 fetal liver myeloid progenitors and LSK were analysed by RT-qPCR. **b**, FACS-sorted HSCs from E14.5 fetal liver and adult bone marrow were analysed by RT-qPCR. **c**, FACS-sorted E14.5 fetal liver and adult bone marrow HSCs were analysed by confocal microscopy. **d**, E14.5 fetal liver TER119<sup>-</sup>CD45<sup>-</sup>CD31<sup>+</sup> endothelial cells (EC), TER119<sup>-</sup>CD45<sup>-</sup>CD31<sup>-</sup>cKit<sup>+</sup>ICAM-1<sup>-</sup> hepatocyte progenitor cells (HP) and TER119<sup>-</sup>CD45<sup>-</sup>CD31<sup>-</sup>cKit<sup>+</sup>ICAM-1<sup>+</sup> mesenchymal cells (MC) were analysed by RT-qPCR (left); negative controls relative to Fig. 1c were analysed by confocal microscopy

(right). **e**, Bone marrow TER119<sup>-</sup>CD45<sup>-</sup>CD31<sup>+</sup>Sca1<sup>+</sup> endothelial cells (Endo), TER119<sup>-</sup>CD45<sup>-</sup>CD31<sup>-</sup>Sca1<sup>+</sup>CD51<sup>+</sup> osteoblasts (Osteo) and TER119<sup>-</sup>CD45<sup>-</sup>CD31<sup>-</sup>Sca1<sup>+</sup>CD51<sup>+</sup> mesenchymal stem cells (MSC) were analysed by RT-qPCR (left) and by confocal microscopy (right). **f**, E14.5 fetal liver and adult bone marrow were analysed by confocal microscopy. Arrows, candidate Lin<sup>-</sup>CD150<sup>-</sup>CD48<sup>-</sup>CD41<sup>-</sup> HSCs, relative to Fig. 1d, e. Figure shows negative controls for GFL staining. **g**, E15.5 gut tissue was analysed by confocal microscopy. White bar, 5  $\mu m$ . Error bars, s.e.m. Housekeeping genes: *Gapdh* and *Hprt1*. \*\*\**P* value for one-way ANOVA lower than 0.001.

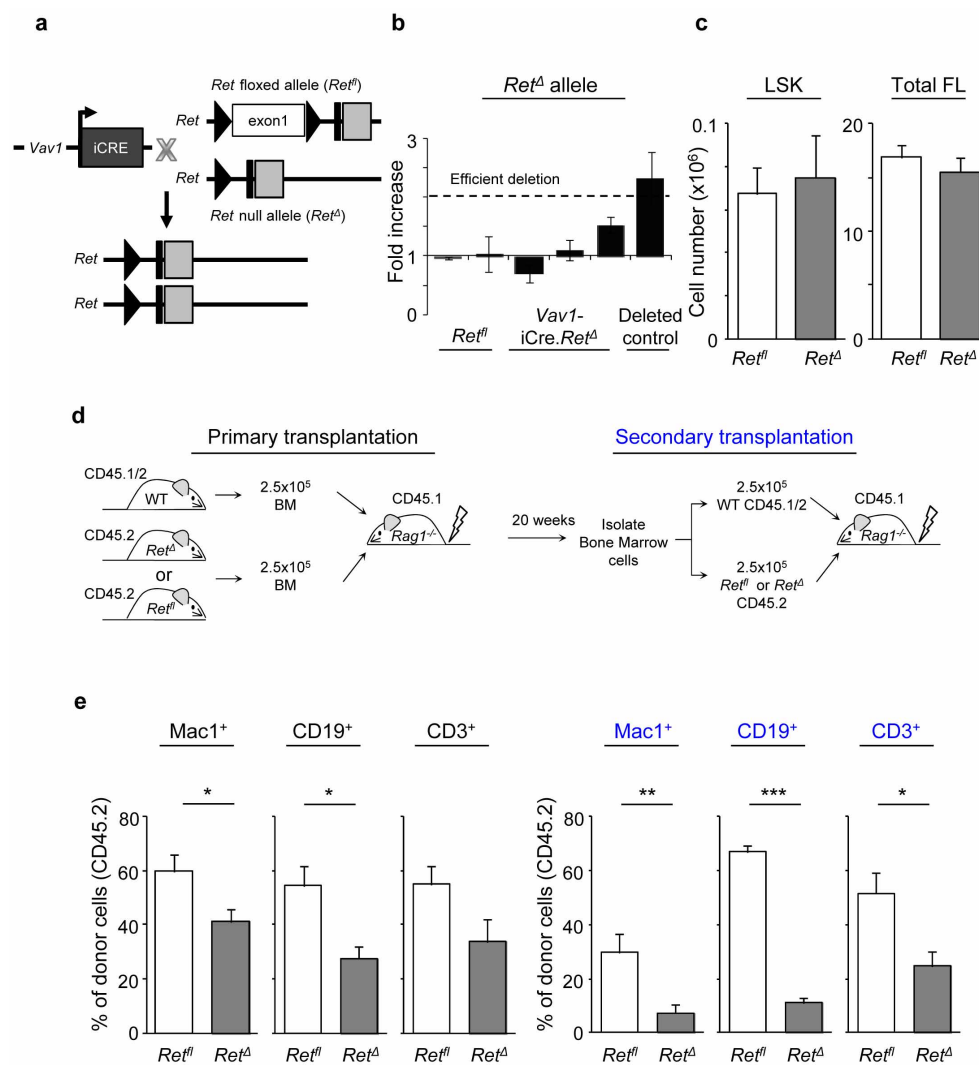


**Extended Data Figure 2 | LSKs are affected by *Ret* deficiency and have reduced reconstitution capacity.** **a**, E14.5 total fetal liver cells. WT  $n = 20$ ; *Ret*<sup>-/-</sup>  $n = 18$ . **b**, Flow cytometry analysis of E14.5 *Ret*<sup>-/-</sup> and WT littermate control LSKs (top) and myeloid progenitors (bottom). **c**, Flow cytometry analysis of Annexin V<sup>+</sup> cells in cultured LSK cells. **d**, Flow cytometry analysis of E14.5 *Ret*<sup>-/-</sup> and WT littermate control HSCs. **e**, Flow cytometry analysis of E14.5 *Ret*<sup>-/-</sup> and WT Lin<sup>-</sup>CD150<sup>+</sup>CD48<sup>-</sup> HSC cells. Mean fluorescence

intensity of cKit was analysed. No statistically significant differences were found. **f**, Lin<sup>-</sup>cKit<sup>+</sup> cells were labelled with CMTMR. Percentage of Lin<sup>-</sup>cKit<sup>+</sup>CMTMR<sup>+</sup> cells in bone marrow and spleen 20 h after injection ( $n = 3$ ). **g**, **h**, Percentage of donor CD45.2 cells in blood cell lineages 16 weeks after primary and secondary (red) transplantation, relative to Fig. 2b, d. **i**, Day 12 CFU-s. WT  $n = 10$ ; *Ret*<sup>-/-</sup>  $n = 10$ . Error bars, s.e.m. \*, \*\*, and \*\*\*,  $P$  values for Student's  $t$ -test lower than 0.05, 0.01 and 0.001 respectively.



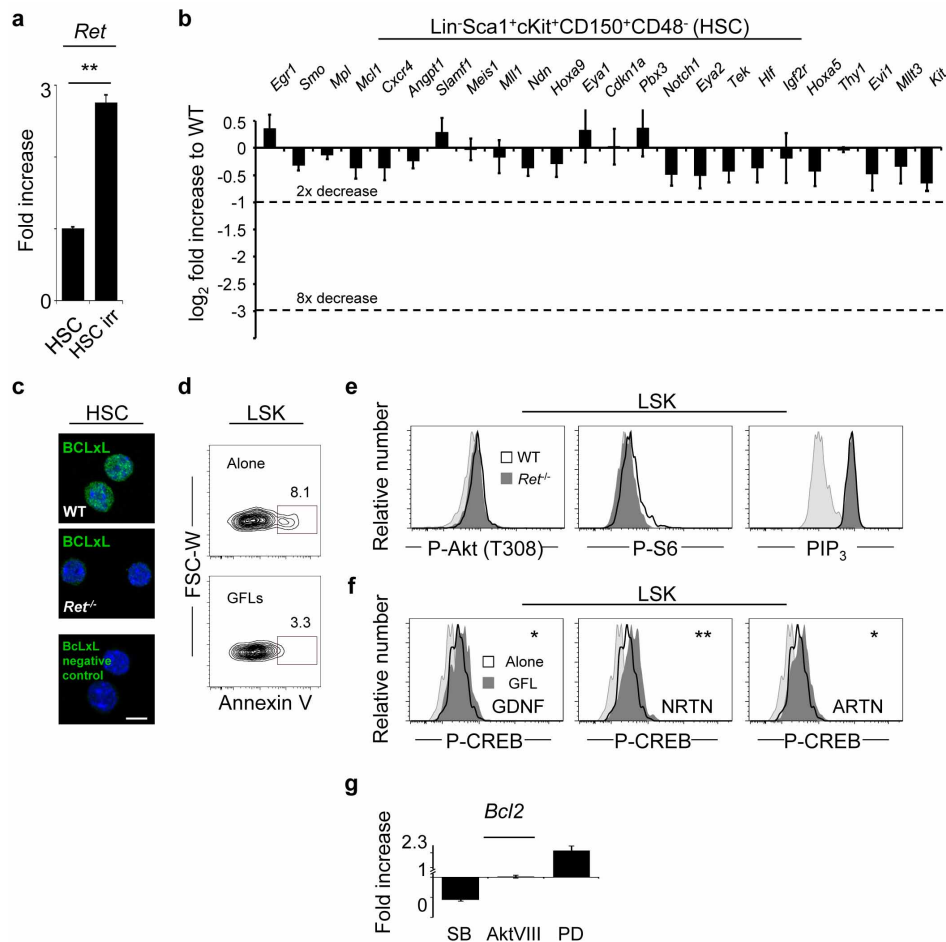
**Extended Data Figure 3 | *Gfra1*-deficient embryos have normal LSK numbers and reconstitution potential.** **a**, Flow cytometry analysis of E14.5 *Gfra1*<sup>-/-</sup> and WT littermate control LSKs (top) and myeloid progenitors (bottom). **b**, Number of LSKs and total fetal liver cells. WT *n* = 9; *Gfra1*<sup>-/-</sup> *n* = 10. **c**, *Gfra1*<sup>-/-</sup> or WT cells were injected with a third-party CD45.1/2 competitor. Percentage of donor CD45.2 cells in blood 16 weeks after transplantation. WT *n* = 5; *Gfra1*<sup>-/-</sup> *n* = 3. **d**, Flow cytometry analysis of bone marrow LSK cells 16 weeks after transplantation. Error bars, s.e.m.



**Extended Data Figure 4 | *Ret* conditional knockout mice and analysis of haematopoietic stem cells.** **a**, *Ret* conditional knockout. Scheme of the targeted *Ret* allele. **b**, E14.5 fetal liver HSCs from *Ret<sup>fl</sup>* littermate controls (two animals) and *Ret* conditional knockout *Vav1-iCre.Ret<sup>Δ</sup>* (three animals) and deleted bone marrow control (last column) were purified by flow cytometry. Efficient deletion of *Ret* in *Vav1-iCre.Ret<sup>Δ</sup>* cells was determined

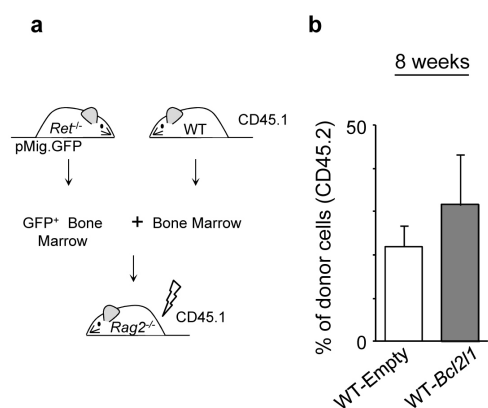
by qPCR as fold increase relative to littermate control cells. **c**, Number of E14.5 LSKs and total fetal liver cells. *Ret<sup>fl</sup>* *n* = 8; *Ret<sup>Δ</sup>* *n* = 8. **d**, Scheme of competitive transplantation with *Ret<sup>Δ</sup>* animals and littermate controls, relative to Fig. 2h. **e**, Percentage of donor CD45.2 cells in blood cell lineages 16 weeks after primary and secondary transplantation. Error bars, s.e.m. \*, \*\* and \*\*\*, *P* values for Student's *t*-test lower than 0.05, 0.01 and 0.001 respectively.



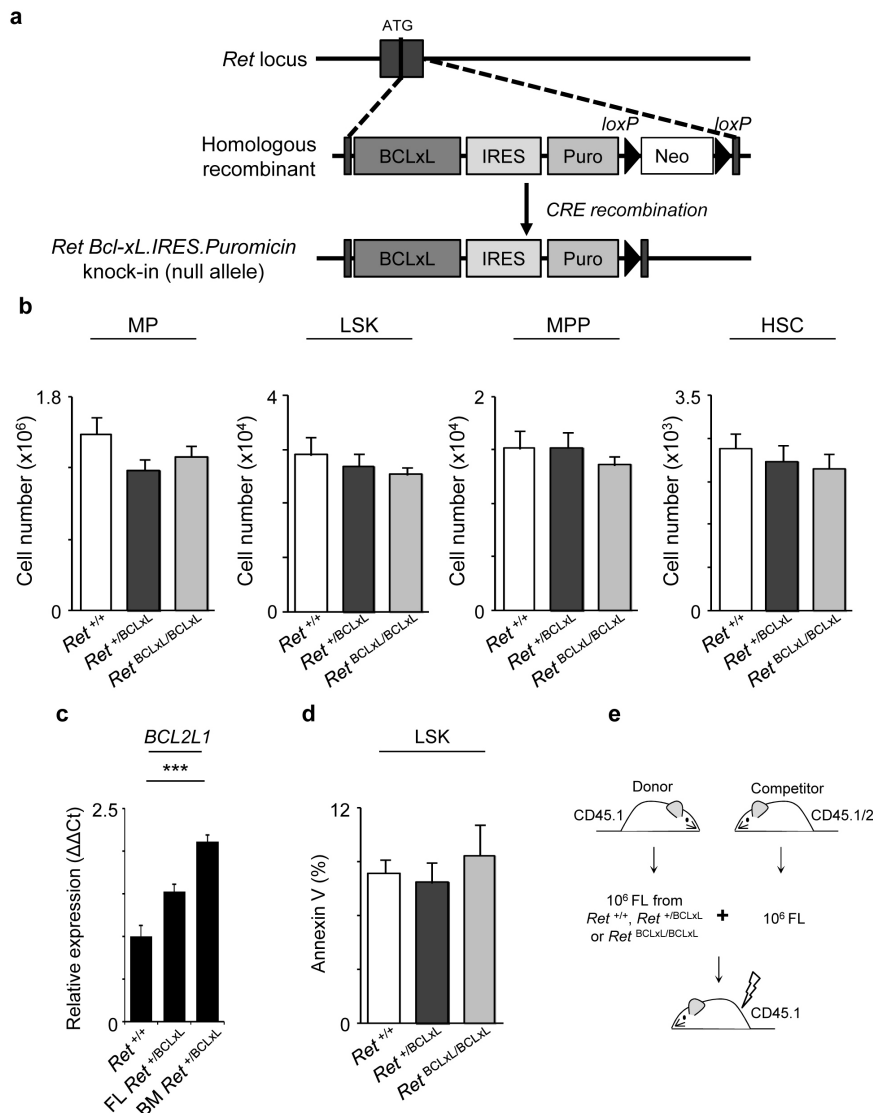


**Extended Data Figure 5 | *Ret* expression increases after haematopoietic stress, and RET signalling increases CREB phosphorylation and cell survival.** **a**, Mice were sublethally irradiated. Irradiation-induced stressed HSCs were purified by flow cytometry at 72 h and analysed by RT-qPCR. **b**, RT-qPCR for fetal liver E14.5 *Ret*<sup>-/-</sup> and WT HSCs ( $n = 3$ ). **c**, Confocal analysis of BCLxL expression in WT and *Ret*<sup>-/-</sup> HSCs. **d**, Flow cytometry analysis of Annexin V<sup>+</sup> cells in cultured LSK cells. **e**, Flow cytometry of E14.5

*Ret*<sup>-/-</sup> and WT littermate controls. P-Akt (T308), P-S6 and PIP<sub>3</sub>: WT  $n = 6$ , *Ret*<sup>-/-</sup>  $n = 6$ . **f**, Flow cytometry analysis of LSK cells in the absence or presence of GDNF, NRTN or ARTN for 1 h ( $n = 6$ ). **g**, *Bcl2* expression in LSK cells upon GFL treatment and with different inhibitors, relative to LSKs treated with GFLs only. Light grey, isotype control. White bar, 5  $\mu$ m. Error bars, s.e.m. \* and \*\*,  $P$  values for Student's  $t$ -test lower than 0.05 and 0.01 respectively.



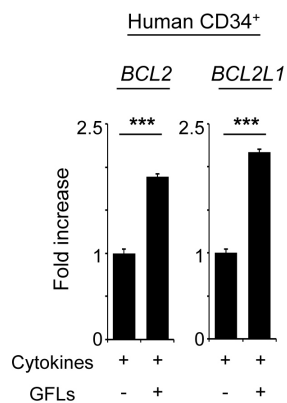
**Extended Data Figure 6 | Rescue of haematopoietic progenitors with *Ret* and its downstream targets.** **a**, Scheme of competitive transplantation, relative to Fig. 4b, c. **b**, Flow cytometry analysis of donor CD45.2 blood cells at 8 weeks upon transplantation of *Bcl2l1*-transduced WT haematopoietic progenitor cells. Error bars, s.e.m.



### Extended Data Figure 7 | Generation and analysis of $Ret^{BCLxL}$ mice.

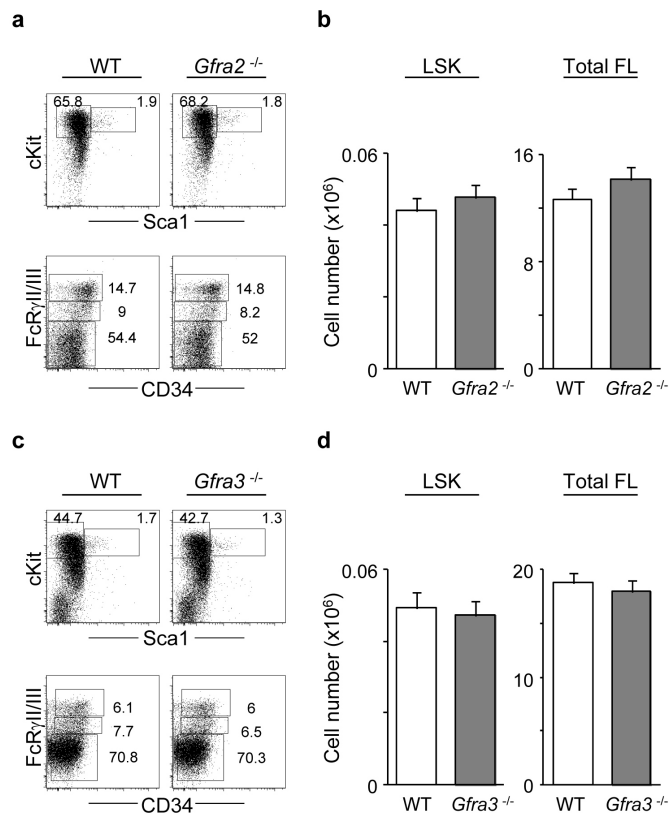
**a**, *Ret* locus was targeted by a construct containing the BCLxL coding sequence, an internal ribosomal entry site (IRES) and a puromycin resistance cassette, followed by a floxed neomycin resistance cassette to aid negative selection. *Ret* Bcl-xL.IRES.Puromycin knock-in mice were obtained by excising of the neomycin cassette. **b**, Number of myeloid progenitors, LSK cells, multipotent progenitor cells and HSCs in E14.5 fetal liver. WT  $n = 7$ ;  $Ret^{+/BCLxL}$   $n = 9$ ;  $Ret^{BCLxL/BCLxL}$   $n = 8$ . **c**, FACS-sorted HSCs from  $Ret^{+/+}$ , E14.5 fetal liver

$Ret^{+/BCLxL}$  and adult bone marrow  $Ret^{+/BCLxL}$  were analysed by RT-qPCR for human *BCL2L1* expression. Housekeeping genes: *Gapdh* and *Hprt1*. **d**, Annexin V<sup>+</sup> cells in cultured E14.5 LSK cells. WT  $n = 6$ ;  $Ret^{+/BCLxL}$   $n = 4$ ;  $Ret^{BCLxL/BCLxL}$   $n = 6$ . **e**, Scheme of competitive transplantation.  $Ret^{BCLxL/BCLxL}$  animals and littermate controls were injected in competition with CD45.1/CD45.2 cells, relative to Fig. 4d, e. Error bars, s.e.m. \*\*\* $P$  value for one-way ANOVA lower than 0.001.

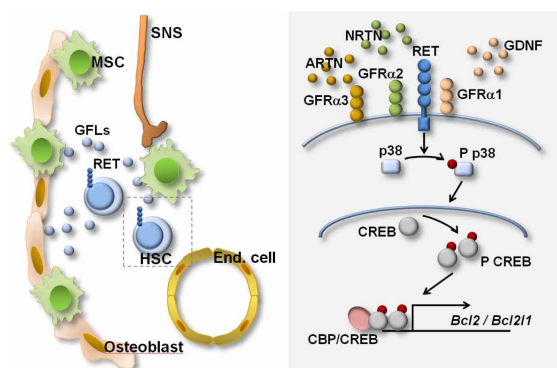


**Extended Data Figure 8 | GFLs increase anti-apoptotic gene expression in human haematopoietic progenitors.** Human cord blood CD34<sup>+</sup> cells were cultured in the presence or absence of GFLs for 4 days and analysed by RT-qPCR. Gene expression relative to cells cultured without GFLs. Error bars, s.e.m. \*\*\**P* values for Student's *t*-test lower than 0.001.





**Extended Data Figure 9 | *Gfra2*- and *Gfra3*-deficient embryos have normal haematopoietic progenitors.** **a**, Flow cytometry analysis of E14.5 *Gfra2*<sup>-/-</sup> and WT littermate control LSKs (top) and myeloid progenitors (bottom). **b**, Number of LSKs and total fetal liver cells. WT *n* = 12; *Gfra2*<sup>-/-</sup> *n* = 11. **c**, Flow cytometry analysis of E14.5 *Gfra3*<sup>-/-</sup> and WT littermate control LSKs (top) and myeloid progenitors (bottom). **d**, Number of LSKs and total fetal liver cells. WT *n* = 11; *Gfra3*<sup>-/-</sup> *n* = 20. Error bars, s.e.m.



**Extended Data Figure 10 | Neuronal growth factors regulate HSC response to physiological demand.** The neurotrophic factors GDNF, NRTN and ARTN are produced by cells in the HSC microenvironment and act directly on HSCs by activation of RET. Highlighted area: RET stimulation results in p38/MAP kinase and CREB activation leading to *Bcl2* and *Bcl2l1* expression. RET signals provide HSCs with survival signals that preserve HSC stemness.

# A long noncoding RNA protects the heart from pathological hypertrophy

Pei Han<sup>1,2</sup>, Wei Li<sup>1,2\*</sup>, Chiou-Hong Lin<sup>2\*</sup>, Jin Yang<sup>1</sup>, Ching Shang<sup>2</sup>, Sylvia T. Nurnberg<sup>2</sup>, Kevin Kai Jin<sup>2</sup>, Weihong Xu<sup>3</sup>, Chieh-Yu Lin<sup>2</sup>, Chien-Jung Lin<sup>2</sup>, Yiqin Xiong<sup>2</sup>, Huan-Chieh Chien<sup>2</sup>, Bin Zhou<sup>4</sup>, Euan Ashley<sup>2</sup>, Daniel Bernstein<sup>5</sup>, Peng-Sheng Chen<sup>1</sup>, Huei-Sheng Vincent Chen<sup>6</sup>, Thomas Quertermous<sup>2</sup> & Ching-Pin Chang<sup>1,7,8</sup>

The role of long noncoding RNA (lncRNA) in adult hearts is unknown; also unclear is how lncRNA modulates nucleosome remodelling. An estimated 70% of mouse genes undergo antisense transcription<sup>1</sup>, including myosin heavy chain 7 (*Myh7*), which encodes molecular motor proteins for heart contraction<sup>2</sup>. Here we identify a cluster of lncRNA transcripts from *Myh7* loci and demonstrate a new lncRNA–chromatin mechanism for heart failure. In mice, these transcripts, which we named myosin heavy-chain-associated RNA transcripts (*Myheart*, or *Mhrt*), are cardiac-specific and abundant in adult hearts. Pathological stress activates the Brg1–Hdac–Parp chromatin repressor complex<sup>3</sup> to inhibit *Mhrt* transcription in the heart. Such stress-induced *Mhrt* repression is essential for cardiomyopathy to develop: restoring *Mhrt* to the pre-stress level protects the heart from hypertrophy and failure. *Mhrt* antagonizes the function of Brg1, a chromatin-remodelling factor that is activated by stress to trigger aberrant gene expression and cardiac myopathy<sup>3</sup>. *Mhrt* prevents Brg1 from recognizing its genomic DNA targets, thus inhibiting chromatin targeting and gene regulation by Brg1. It does so by binding to the helicase domain of Brg1, a domain that is crucial for tethering Brg1 to chromatinized DNA targets. Brg1 helicase has dual nucleic-acid-binding specificities: it is capable of binding lncRNA (*Mhrt*) and chromatinized—but not naked—DNA. This dual-binding feature of helicase enables a competitive inhibition mechanism by which *Mhrt* sequesters Brg1 from its genomic DNA targets to prevent chromatin remodelling. A *Mhrt*–Brg1 feedback circuit is thus crucial for heart function. Human *MHRT* also originates from *MYH7* loci and is repressed in various types of myopathic hearts, suggesting a conserved lncRNA mechanism in human cardiomyopathy. Our studies identify a cardioprotective lncRNA, define a new targeting mechanism for ATP-dependent chromatin-remodelling factors, and establish a new paradigm for lncRNA–chromatin interaction.

By 5' and 3' rapid amplification of complementary DNA ends, we discovered an alternative splicing of *Myh7* antisense transcription into a cluster of RNAs of 709 to 1,147 nucleotides (*Mhrt* RNAs), containing partial sequences of *Myh7* introns and exons (Fig. 1a and Supplementary Note). *Mhrt* RNAs were cardiac-specific (Fig. 1b), present at low levels in fetal hearts, with increasing abundance as the hearts matured and *Myh6/Myh7* ratio increased (Fig. 1c). RNA *in situ* analysis showed that *Mhrt* RNAs resided in the myocardium but not endocardium or epicardium (Fig. 1d and Extended Data Fig. 1a). Quantification of nuclear/cytoplasmic RNA in heart extracts revealed that *Mhrt* transcripts were primarily nuclear RNAs (Fig. 1e). Coding substitution frequencies<sup>4,5</sup> of *Mhrt* RNAs predicted a negative/low protein-coding potential, *in vitro* translation of *Mhrt* RNAs yielded no proteins, and ribosome profiling<sup>6</sup> revealed

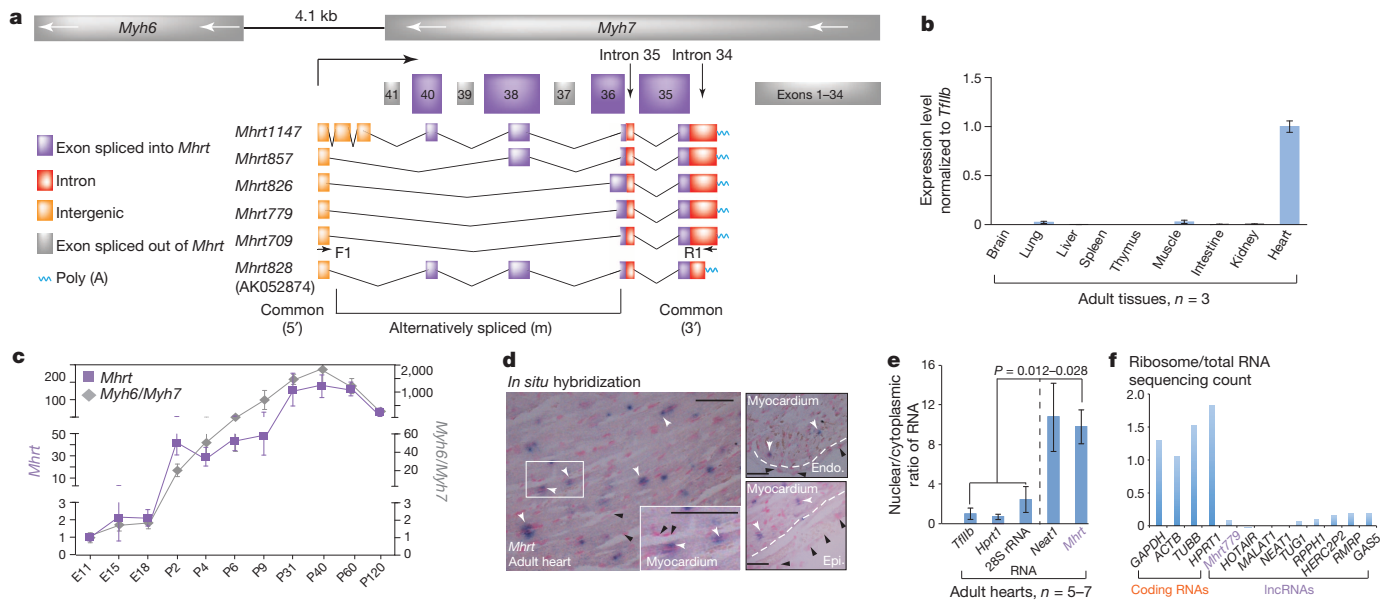
no/minimal ribosomes on *Mhrt* (Fig. 1f, Extended Data Fig. 1b–f and Supplementary Note). Consequently, *Mhrt* RNAs are non-coding RNAs in cardiomyocyte nuclei.

*Mhrt* RNAs were downregulated by 46–68% in hearts pressure-overloaded by transaortic constriction (TAC)<sup>3</sup>, beginning by 2 days and lasting for ≥42 days after TAC (Fig. 2a). Such *Mhrt* reduction coincided with the TAC-induced *Myh6* to *Myh7* isoform switch characteristic of cardiomyopathy<sup>7–9</sup> (Extended Data Fig. 2a). To define *Mhrt* function, we focused on *Mhrt779*, the most abundant *Mhrt* species, with 779 nucleotides (Fig. 2b, c and Extended Data Fig. 2b–e). We generated a transgenic mouse line to restore *Mhrt779* level in stressed hearts. This transgenic line, driven by tetracycline response element (*Tre-Mhrt779*), was crossed to a cardiac-specific driver line (*Tnnt2-rtTA*)<sup>3</sup> that employs troponin promoter (*Tnnt2*) to direct expression of reverse tetracycline-dependent transactivator (rtTA). The resulting *Tnnt2-rtTA;Tre-Mhrt779* line (abbreviated as *Tg779*) enabled the use of doxycycline to induce *Mhrt779* expression in cardiomyocytes. Within 7–14 days of doxycycline treatment, *Mhrt779* increased by ~1.5-fold in left ventricles of *Tg779* mice; this offset *Mhrt779* suppression in TAC-stressed hearts to maintain *Mhrt779* at the pre-stress level (Fig. 2d). Six weeks after TAC, doxycycline-treated control mice (*Tre-Mhrt779*, *Tnnt2-rtTA* or wild type) developed severe cardiac hypertrophy and fibrosis with left ventricular dilatation and reduced fractional shortening. Conversely, doxycycline-treated *Tg779* hearts—with *Mhrt779* maintained at the pre-stress level—developed much less pathology, with a 45.7% reduction in the ventricle/body-weight ratio (Fig. 2e) and a 61.3% reduction in cardiomyocyte size (Fig. 2f and Extended Data Fig. 3a), minimal/absent cardiac fibrosis (Fig. 2g), a 45.5% improvement of fractional shortening (Fig. 2h and Extended Data Fig. 3b), normalized left ventricular size (Fig. 2i), and reduced pathological changes of *Anf* (also known as *Nppa*), *Bnp* (also known as *Nppb*), *Serca2* (also known as *Atp2a2*), *Tgfb1* and *Opn* (also known as *Spp1*) expression<sup>10–13</sup> (Extended Data Figs 3c and 6e). To further test the cardioprotective effects of *Mhrt*, we induced *Mhrt779* after 1–2 weeks of TAC when hypertrophy had begun. This approach reduced hypertrophy by 23% and improved fractional shortening by 33% in 8 weeks after TAC (Extended Data Fig. 3d–f). The efficacy of late *Mhrt779* introduction suggests that a sustained repression of *Mhrt* in stressed hearts is essential for continued decline of cardiac function.

To study *Mhrt* regulation, we examined the 5' upstream region of the *Mhrt* genomic site (–2329 to +143) (Extended Data Fig. 4a) for signatures of a lncRNA promoter: RNA polymerase II (Pol II), histone H3 trimethylated lysine 4 (H3K4me3) and histone H3 trimethylated lysine 36 (H3K36me3)<sup>4,14,15</sup>. By chromatin immunoprecipitation (ChIP) of left ventricles, we found that this putative promoter contained four evolutionarily conserved elements (a1 to a4)<sup>3</sup> that were enriched with Pol II

<sup>1</sup>Krannert Institute of Cardiology and Division of Cardiology, Department of Medicine, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA. <sup>2</sup>Division of Cardiovascular Medicine, Cardiovascular Institute, Stanford University School of Medicine, Stanford, California 94305, USA. <sup>3</sup>Stanford Genome Technology Center, Stanford University School of Medicine, Stanford, California 94305, USA. <sup>4</sup>Department of Genetics, Pediatrics, and Medicine (Cardiology), Albert Einstein College of Medicine of Yeshiva University, 1301 Morris Park Avenue, Price Center 420, Bronx, New York 10461, USA. <sup>5</sup>Department of Pediatrics, Stanford University School of Medicine, Stanford, California 94305, USA. <sup>6</sup>Del E. Webb Neuroscience, Aging & Stem Cell Research Center, Sanford/Burnham Medical Research Institute, La Jolla, California 92037, USA. <sup>7</sup>Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA. <sup>8</sup>Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA.

\*These authors contributed equally to this work.



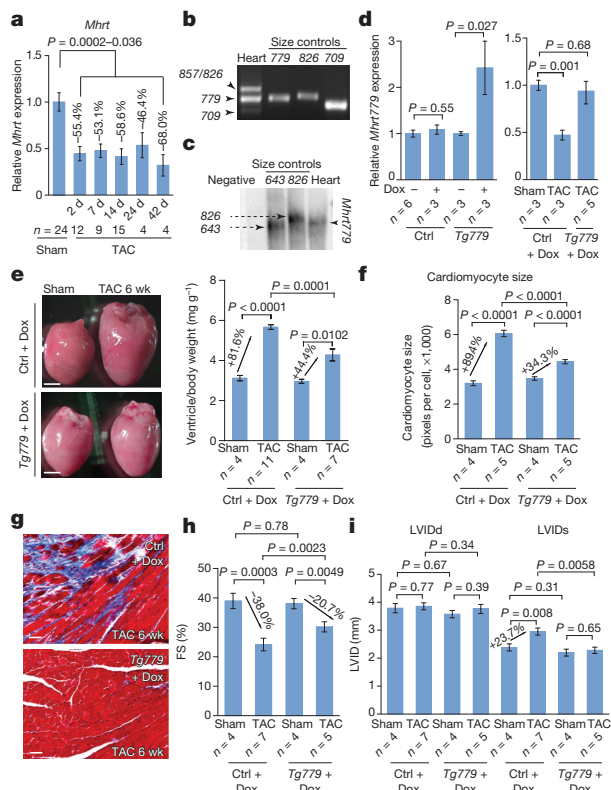
**Figure 1 | Profile of the noncoding RNA *Mhrt*.** **a**, Schematic illustration of *Mhrt* RNAs originating from the intergenic region between *Myh6* and *Myh7* and transcribed into *Myh7*. *Myh7* exons and introns are indicated. m, mid region of the RNAs. F1 and R1, targeting the 5' and 3' *Mhrt* common sequences, are the primers used for subsequent polymerase chain reaction (PCR). **b**, Quantitative PCR with reverse transcription (RT-qPCR) of *Mhrt* RNAs using primers targeting common regions of *Mhrt* RNAs in tissues from 2-month-old mice. **c**, RT-qPCR of *Mhrt*, *Myh6* and *Myh7* in mouse hearts at different ages. *Mhrt* and *Myh6/Myh7* ratio of embryonic day (E)11 hearts are set as 1. **d**, RNA *in situ* analysis of *Mhrt* (blue) in adult hearts.

The RNA probe targets all *Mhrt* species. Red: nuclear fast red. White arrowheads indicate myocardial nuclei. Black arrowheads indicate nuclei of endothelial, endocardial or epicardial cells. Dashed lines demarcate the myocardium from endocardium (Endo.) or from epicardium (Epi.). Scale bars = 50  $\mu$ m. **e**, RT-qPCR of nuclear/cytoplasmic RNA in adult hearts. *Tlfb* (also known as *Gtf2b*), *Hprt1* and 28S rRNA are primarily cytoplasmic RNAs; *Neat1*, nuclear lncRNA. *Tlfb* ratio is set as 1. **f**, Ribosome profiling: ribosome density on coding RNAs and lncRNAs. *P* values: Student's *t*-test. Error bars show standard error of the mean (s.e.m.).

(a1 to a4), H3K4me3 (a1 and a4) and H3K36me3 (refs 14, 16–18) (a1 and a3/a4) (Extended Data Fig. 4a–d). Conversely, no Pol II, H3K4me3 or H3K36me3 enrichment was found in control *Shh* and *Vegfa* promoters or in thymus and lungs that did not express *Mhrt* RNAs (Extended

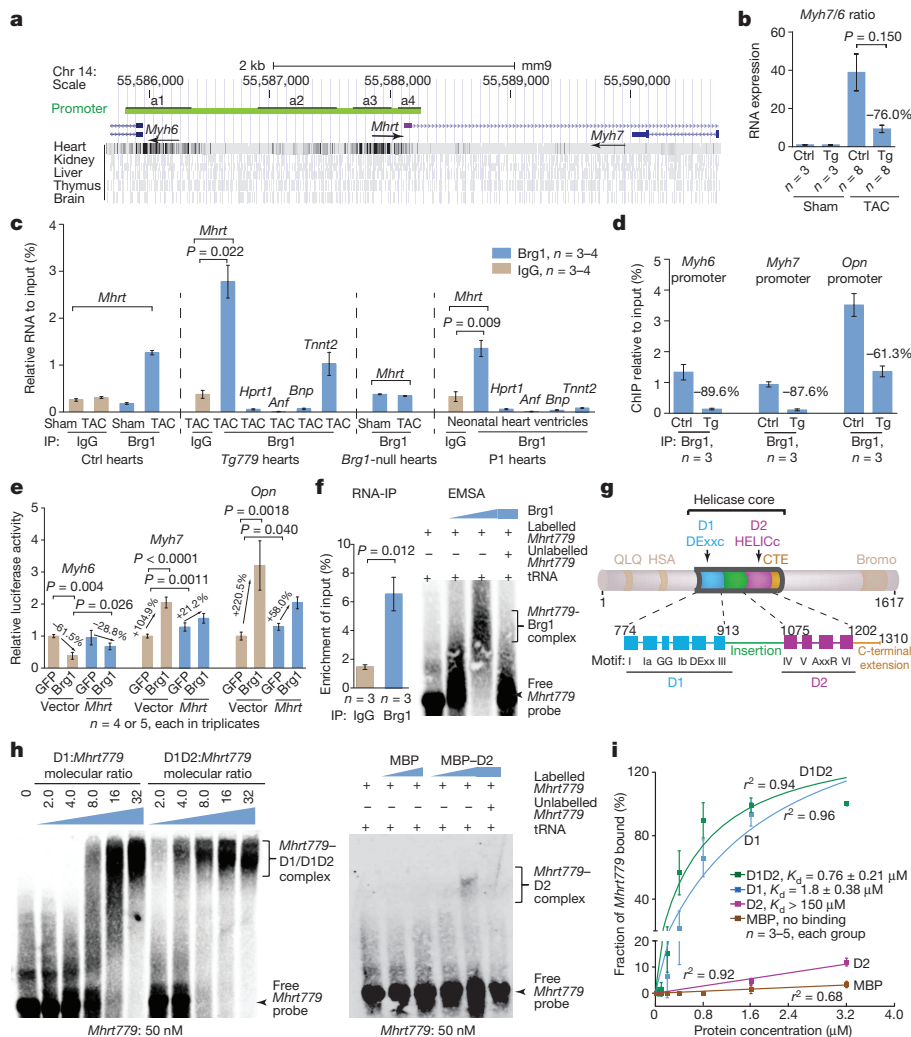
Data Fig. 4b–d). These results reveal an active, cardiac-specific lncRNA promoter controlling *Mhrt* expression.

We then asked how *Mhrt* was repressed in stressed hearts. We postulated that cardiac stress activated Brg1, leading it to occupy the a1–a4 promoter and to repress *Myh6* (ref. 3) and *Mhrt* in opposite transcription directions (Extended Data Fig. 4a). Indeed, *Mhrt* repression required Brg1: TAC suppressed *Mhrt* RNAs in control but not *Brg1*-null hearts (*Tnnt2-rtTA; Tre-Cre; Brg1<sup>fl/fl</sup>*)<sup>3</sup> (Extended Data Fig. 4e). To test Brg1 activity on the *Mhrt* promoter, we cloned the a1–a4 promoter in the *Mhrt* transcription direction (–2329 to +143) into an episomal luciferase reporter, pREP4, that allows promoter chromatinization<sup>19</sup>. Brg1 was then transfected into Brg1-deficient SW13 cells<sup>20</sup> to reconstitute the Brg1/BAF complex for reporter assays. Brg1 transfection caused a ~50% reduction of *Mhrt* promoter activity ( $P < 0.0001$ ), and such *Mhrt* repression was virtually abolished by Hdac inhibition with trichostatin-A or Parp inhibition



**Figure 2 | *Mhrt* inhibits cardiac hypertrophy and failure.** **a**, Quantification of cardiac *Mhrt* RNAs 2–42 days (d) after TAC operation. **b**, RT-PCR of *Mhrt* RNAs in adult heart ventricles. Primers (F1 and R1; Fig. 1a) target *Mhrt* common regions. Size controls 779, 826 and 709 are PCR products of recombinant *Mhrt779*, *Mhrt826* and *Mhrt709*, respectively. **c**, Northern blot of *Mhrt* RNAs in adult heart ventricles. The probe targets common regions of *Mhrt* RNAs. Negative: control RNA from 293T cells. Size control 826 is recombinant *Mhrt826*; 643 (not a distinct *Mhrt* species) contains the 5' common region of *Mhrt*. **d**, Quantification of *Mhrt779* expression in control or *Tg779* mice with or without doxycycline (Dox) or TAC operation. *Mhrt779*-specific PCR primers were used. Ctrl, control mice. **e**, Ventricle/body-weight ratio of hearts 6 weeks (wk) after TAC. Scale bars = 1 mm. **f**, Quantification of cardiomyocyte size in control and *Tg779* mice 6 weeks after TAC by wheat-germ agglutinin staining. **g**, Trichrome staining in control and *Tg779* hearts 6 weeks after TAC. Red indicates cardiomyocytes; blue indicates fibrosis. Scale bars = 20  $\mu$ m. **h**, **i**, Echocardiographic measurement of left ventricular fractional shortening (FS; **h**) and internal dimensions at end-diastole (LVIDd) and end-systole (LVIDs) (**i**) 6 weeks after TAC. *P* values: Student's *t*-test. Error bars show s.e.m.





**Figure 3 | *Mhrt* complexes with Brg1 through the helicase domain.** **a**, DNaseI digital footprinting of *Myh6/Mhrt* promoter loci from ENCODE. *Myh6* and *Mhrt* are transcribed in opposite directions as indicated by arrows. Bars represent DNA fragments protected from DNaseI digestion. Black boxes (a1–a4) refer to promoter regions with high sequence homology (Extended Data Fig. 4a). **b**, Quantification of *Myh7/Myh6* ratio in control (Ctrl) and *Tg779* (Tg) hearts 2 weeks after TAC. **c**, RNA-immunoprecipitation (IP) of *Mhrt*–Brg1 in ventricles from control hearts (Ctrl) with sham/TAC operation; *Tg779* hearts after TAC; *Brg1*-null (*Tnnt2-rtTA;Tre-Cre;Brg1<sup>fl/fl</sup>*) hearts after TAC; and P1 hearts. **d**, ChIP analysis of Brg1 in control (Ctrl) and *Tg779* hearts 2 weeks after TAC. **e**, Luciferase reporter assay of *Myh6* and *Myh7* promoters in neonatal rat cardiomyocytes. *Mhrt*, pAdd2-*Mhrt779*; Vector, pAdd2 empty vector. **f**, RNA-IP and EMSA of recombinant Brg1 proteins and *in vitro* transcribed *Mhrt779*. Biotin-labelled *Mhrt779*; 50 nM; unlabelled *Mhrt779*; 500 nM. **g**, Schematic of mouse Brg1 protein. The helicase core includes the DEXx-c and HELIC-c domain. **h**, EMSA of *Mhrt779* and Brg1 helicase. **i**, Binding affinity of *Mhrt779* for MBP-tagged D1D2 determined by EMSA. Data are from multiple independent measurements. Nonlinear regression curves were generated by GraphPad Prism. *P* values: Student's *t*-test. Error bars show s.e.m.

with PJ-34 (ref. 21) (Extended Data Fig. 4f), indicating a cooperative repressor function between Brg1, Hdac and Parp. ChIP verified that the *Mhrt* promoter (a1–a4) was occupied by Brg1, Hdac2/9 and Parp1 in stressed hearts<sup>3</sup> and in the pREP4 reporter episome (Extended Data Fig. 4g). These findings indicate that *Mhrt* is repressed by the stress-induced Brg1–Hdac–Parp complex<sup>3</sup> through the a1–a4 promoter.

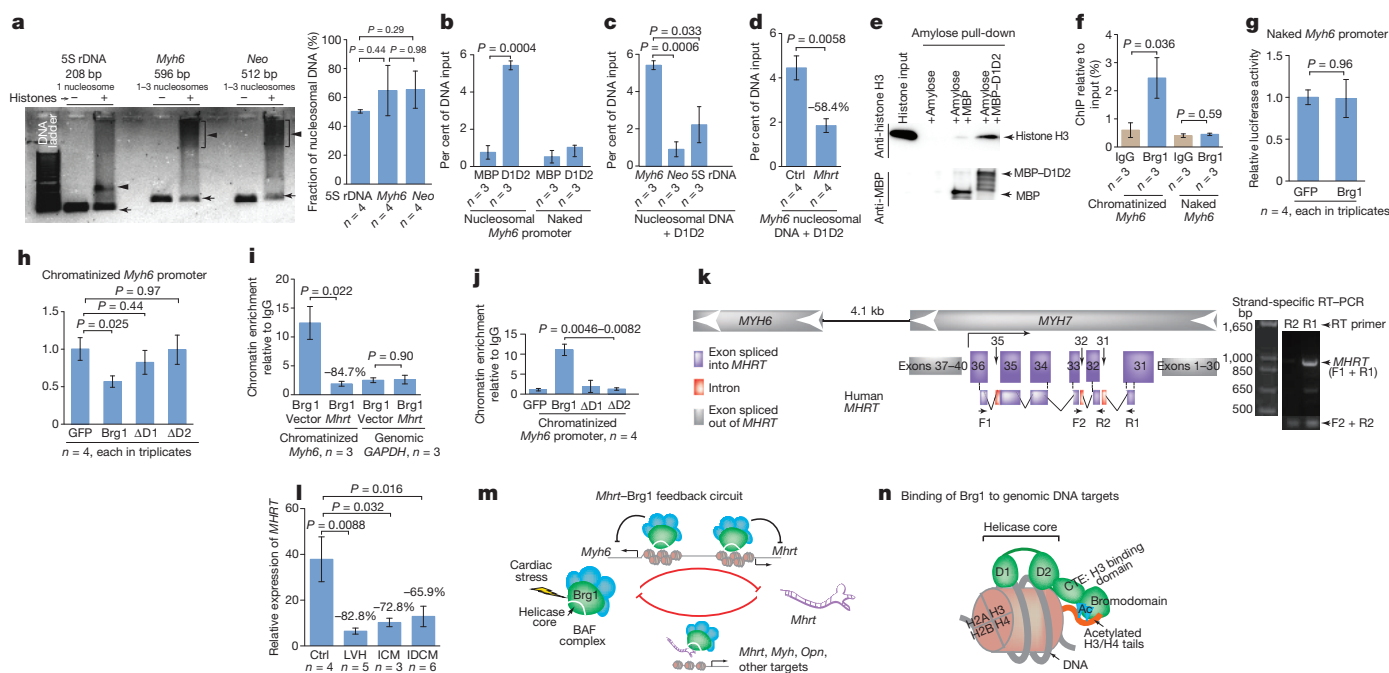
Because *Myh6* and *Mhrt* were both regulated by the a1–a4 promoter, we hypothesized that a1–a4 contained two elements to regulate *Myh6* and *Mhrt*—with the a1 element controlling *Myh6* and the a3/4 element controlling *Mhrt* (Extended Data Fig. 4a). On a1 and a3/4 (but not a2), we found cardiac-specific enrichment of Brg1 (ref. 3), H3K4me3 and H3K36me3 (Extended Data Fig. 4c–d), and DNaseI genomic footprints (Fig. 3a)<sup>22</sup>. To test a3/4 for *Mhrt* regulation, we conducted deletional analysis of the a1–a4 promoter in the *Mhrt* transcription direction. In reporter assays, a3/4 was necessary and sufficient for *Mhrt* promoter activity and for Brg1-dependent *Mhrt* repression, whereas a1 was not essential for either (Extended Data Fig. 4h). Conversely, a1 is necessary and sufficient for Brg1 to repress the *Myh6* promoter<sup>3</sup>, but a3/4 is not required<sup>3</sup>. Therefore, a1 and a3/4 are two functionally distinct elements for Brg1 to separately control *Myh6* and *Mhrt*.

In stressed hearts, Brg1 represses *Myh6* and activates *Myh7* (ref. 3), causing a pathological switch of *Myh6/7* expression, contributing to cardiomyopathy<sup>23</sup>. This stress/Brg1-dependent *Myh* switch was largely eliminated by *Mhrt779* (Fig. 3b), and the inhibition of the *Myh* switch by *Mhrt* did not involve RNA–RNA sequence interference between *Mhrt* and *Myh* (Extended Data Fig. 5a–j and Supplementary Note). Instead, it required a physical interaction between *Mhrt* RNA and Brg1. RNA

immunoprecipitation of TAC-stressed adult hearts or Brg1-expressing neonatal hearts showed that Brg1 co-immunoprecipitated with *Mhrt779* but not control RNAs, and that *Mhrt779* complexed with Brg1 but not with the polycomb proteins Ezh2 or Suz12 (Fig. 3c and Extended Data Fig. 6a, b). The Brg1–*Mhrt* complex was minimal in unstressed adult hearts with low Brg1 (ref. 3) or in stressed *Brg1*-null hearts (*Tnnt2-rtTA;Tre-Cre;Brg1<sup>fl/fl</sup>*)<sup>3</sup> (Fig. 3c and Supplementary Note). These results suggest that *Mhrt* binds to Brg1 to influence its gene regulation.

We then tested how *Mhrt* regulated Brg1 activity on its *in vivo* target genes, including *Myh6* (ref. 3), *Myh7* (ref. 3) and *Opn* (osteopontin, critical for cardiac fibrosis<sup>12</sup>) (Extended Data Fig. 6c–e and Supplementary Note). In doxycycline-treated, TAC-stressed *Tg779* hearts, *Mhrt779*—without affecting the Brg1 messenger RNA/protein level (Extended Data Fig. 7a–f)—reduced Brg1 occupancy on *Myh6*, *Myh7* and *Opn* promoters by 60–90% (Fig. 3d), causing a 56–76% loss of Brg1-controlled *Myh* switch and *Opn* activation (Fig. 3b and Extended Data Figs 6e, 7g). We then used primary rat ventricular cardiomyocytes to conduct reporter assays. In these cells, as observed *in vivo*, Brg1 repressed *Myh6* and activated *Myh7* and *Opn* promoters; *Mhrt779* reduced Brg1 activity on these promoters by 54–80% (Fig. 3e). Accordingly, *Mhrt* prevents Brg1 from binding to its genomic targets to control gene expression.

How Brg1 or ATP-dependent chromatin remodellers recognize their target promoters is an important but not fully understood issue in chromatin biology. Biochemically, recombinant Brg1 proteins and *in vitro* transcribed *Mhrt779* could directly co-immunoprecipitate without involving other factors (Fig. 3f). An electrical mobility shift assay (EMSA) showed that Brg1 shifted biotin-labelled *Mhrt779* to form a low mobility



**Figure 4 | *Mhrt* inhibits chromatin targeting and gene regulation by Brg1.**

**a**, Gel electrophoresis and quantification of nucleosomal 5S rDNA, the *Myh6* promoter and *Neo* DNA. Arrowheads indicate the DNA–histone complex; arrows indicate naked DNA. Nucleosome assembly efficiency is defined as the fraction of DNA bound to histones (arrowheads). **b–d**, Quantification of amylose pull-down of MBP–D1D2 (D1D2) with nucleosomal and naked *Myh6* promoter DNA (**b**), with nucleosomal *Myh6* promoter, *Neo* and 5S rDNA (**c**), or with nucleosomal *Myh6* promoter in the presence of *Mhrt779* (**d**). **e**, Amylose pull-down of MBP–D1D2 and histone H3. Anti-histone H3 and anti-MBP antibodies were used for western blot analysis. **f**, ChIP analysis of Brg1 on chromatinized and naked *Myh6* promoter in rat ventricular cardiomyocytes. GFP, green fluorescent protein control. **g**, **h**, Luciferase reporter activity of Brg1 on naked *Myh6* promoter (**g**) or of helicase-deficient Brg1 on chromatinized *Myh6* promoter (**h**) in rat ventricular cardiomyocytes.  $\Delta$ D1, Brg1 lacking amino acid 774–913;  $\Delta$ D2, Brg1 lacking 1086–1246. ChIP: H-10 antibody recognizing N terminus, non-disrupted region of Brg1. **i**, **j**, ChIP

analysis in SW13 cells of chromatinized *Myh6* promoter in the presence of *Mhrt779* (**i**) or helicase-deficient Brg1 (**j**). *Mhrt*, pAdd2-*Mhrt779*; Vector, pAdd2 empty vector. **k**, Schematic illustration and PCR of human *MHRT*. *MHRT* originates from *MYH7* and is transcribed into *MYH7*. *MYH7* exons and introns are indicated. R1 and R2 are strand-specific primers; F1 and R1 target *MHRT* and *MYH7*; F2 and R2 are specific for *MHRT*. **l**, Quantification of *MHRT* in human heart tissues. Ctrl, control; ICM, ischaemic cardiomyopathy; IDCM, idiopathic dilated cardiomyopathy; LVH, left ventricular hypertrophy. **m**, Working model of a Brg1–*Mhrt* feedback circuit in the heart. Brg1 represses *Mhrt* transcription, whereas *Mhrt* prevents Brg1 from recognizing its chromatin targets. Brg1 functions through two distinct promoter elements to bidirectionally repress *Myh6* and *Mhrt* expression. **n**, Molecular model of how Brg1 binds to its genomic DNA targets. Brg1 helicase (D1D2) binds chromatinized DNA, C-terminal extension (CTE) binds histone H3, and bromodomain binds acetylated histone H3 or H4. *P* values: Student's *t*-test. Error bars show s.e.m.

protein–RNA complex that was competitively disrupted by unlabelled *Mhrt779* (Fig. 3f). Brg1, which belongs to the SWI/SNF family of chromatin-remodelling factors, contains a helicase/ATPase core that is split by an insertion into two RecA-like domains: DEAD-like helicase superfamily C-terminal domain, D1 (DExx-c) and helicase superfamily C-terminal domain, D2 (HELIC-c)<sup>24,25</sup>, with signature motifs of DEAD-box, superfamily 2 RNA helicase<sup>25,26</sup> (Fig. 3g and Extended Data Fig. 8). SWI/SNF proteins although conserved with RNA helicases, were observed to bind DNA<sup>27</sup> and mediate DNA structural changes and repair<sup>19</sup>. The binding properties of Brg1 remained undefined. To test whether *Mhrt* could bind to Brg1 helicase, we generated maltose-binding protein (MBP)-tagged recombinant proteins that contained the Brg1 DExx-c domain (MBP–D1, amino acids 774–913), the HELIC-c domain with C-terminus extension (MBP–D2, 1086–1310), or the entire helicase (MBP–D1D2, 774–1310) (Extended Data Fig. 9a). D1D2 showed the highest *Mhrt* binding affinity (dissociation constant ( $K_d$ ) = 0.76  $\mu$ M); D1 showed moderate affinity ( $K_d$  = 1.8  $\mu$ M); D2 modest affinity ( $K_d$  > 150  $\mu$ M); and MBP did not bind at all (Fig. 3h, i). Therefore, Brg1 helicase binds *Mhrt* with high affinity.

Contrary to its potent RNA binding, Brg1 helicase showed no detectable binding to the naked DNA of the *Myh6* promoter (596 bp, –426 to +170, critical for the control of *Myh6* by Brg1 (ref. 3)) (Extended Data Fig. 9b). To test whether Brg1 helicase could bind chromatinized DNA, we generated nucleosomal DNA *in vitro* by assembling histone octamers (histones H2A, H2B, H3 and H4)<sup>28</sup> on *Myh6* promoter DNA, as well as

on control neomycin phosphotransferase gene (*Neo*) and 5S ribosomal (r)DNA (5S rDNA). We achieved 50–65% efficiency of nucleosome assembly, comparable between *Myh6*, *Neo* and 5S rDNA (Fig. 4a). Because the large nucleosome size precluded a clear EMSA resolution, we used amylose to pull down MBP-tagged D1D2 proteins. We found that D1D2 pulled down nucleosomal *Myh6* promoter DNA but not the naked one (Fig. 4b). The pull-down efficiency of nucleosomal *Myh6* was ~3–6-fold that of *Neo* or 5S rDNA (Fig. 4c), and *Mhrt779* was capable of disrupting D1D2–*Myh6* pull-down (Fig. 4d). Although D1D2 bound to histone H3 (Fig. 4e), histone binding was insufficient to anchor D1D2 to nucleosomal DNA, as D1D2 bound poorly to nucleosomal *Neo* and 5S rDNA that also contained histones (Fig. 4c). Therefore, chromatinized DNA targets are biochemically recognized by Brg1 helicase, and this process is inhibited by *Mhrt*.

To test the ability of Brg1 to distinguish chromatinized from naked DNA promoters in cells, we cloned *Myh6* promoter into the luciferase reporter plasmid pREP4 (allowing promoter chromatinization<sup>19</sup>) and pGL3 (containing naked, non-chromatinized promoter). In rat ventricular cardiomyocytes and SW13 cells, ChIP and luciferase analyses showed that Brg1 bound and repressed chromatinized but not naked *Myh6* promoter (Fig. 4f, g and Extended Data Fig. 9c, d). However, without D1/D2 domain or in the presence of *Mhrt*, Brg1 was unable to bind or repress chromatinized *Myh6* promoter (Fig. 4h–j and Extended Data Fig. 9e), indicating the necessity of D1D2 for the interaction between Brg1, chromatin and *Mhrt*. Consistently, all our genetic, biochemical and cellular

studies show that Brg1 requires the helicase domain to bind to chromatinized DNA targets, and *Mhrt* seizes the helicase to disrupt Brg1–chromatin binding.

We then asked how Brg1 surpassed its basal suppression by *Mhrt* to control *Myh*, *Mhrt*, *Opn*, or other genes to trigger cardiomyopathy (Supplementary Note). Amylose pull-down experiments showed that Brg1 dose-dependently escaped from *Mhrt* inhibition to occupy *Mhrt* promoter (Extended Data Fig. 10). Brg1 protein, which increases under stress conditions<sup>3</sup>, could therefore outrun *Mhrt* and gain control over the *Mhrt* promoter to repress *Mhrt* expression and tip the balance towards Brg1. Contrary to the endogenous *Mhrt* that was repressible by Brg1, the *Mhrt* transgene (*Tg779*)—driven by *Tnnt2/Tre* promoters—was not subject to repression by Brg1 and was thus able to keep *Mhrt* at pre-stress levels to inhibit Brg1 and reduce hypertrophy. This further demonstrates the necessity of *Mhrt* repression for myopathy to develop.

Human *MYH7* loci encoded RNA that resembled *Mhrt* in primary sequence and secondary structure, as predicted by minimal free energy<sup>29</sup> (Fig. 4k and Extended Data Fig. 11a, b). Human *MHRT* was also repressed in stressed hearts, with 82.8%, 72.8% and 65.9% reduction of *MHRT* in hypertrophic, ischaemic or idiopathic cardiomyopathy tissues, respectively (Fig. 4l and Extended Data Fig. 11c). This suggests a conserved *MHRT* mechanism of human cardiomyopathy.

*Mhrt* is the first example, to our knowledge, of a lncRNA that inhibits myopathy and chromatin remodellers. Reciprocal *Mhrt*–Brg1 inhibition constitutes a feedback circuit critical for maintaining cardiac function (Fig. 4m). The helicase core of Brg1, combined with the histone-binding domains of the Brg1/BAF complex, adds a new layer of specificity control to Brg1/BAF targeting and chromatin remodelling (Fig. 4n). The *Mhrt*–helicase interaction also exemplifies a new mechanism by which lncRNA controls chromatin structure. To further elucidate chromatin regulation, it will be essential to define helicase domain function in all ATP-dependent chromatin-remodelling factors and to identify new members of lncRNA that act through this domain to control chromatin. The cardioprotective *Mhrt* may have translational value, given that RNA can be chemically modified and delivered as a therapeutic drug. This aspect of lncRNA–chromatin regulation may also inspire new therapies for human disease.

## METHODS SUMMARY

*Tg779* mouse generation, rapid amplification of cDNA ends (RACE), RNA *in situ* hybridization, RT–qPCR, codon substitution frequencies (CSF), echocardiography, northern blot, EMSA, ChIP, RNA immunoprecipitation, reporter assay, nucleosome assembly, and the amylose pull-down assay were performed as described<sup>13,4,28</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 March 2013; accepted 17 June 2014.

Published online 10 August; corrected online 1 October 2014 (see full-text HTML version for details).

1. RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group) and the FANTOM Consortium. Antisense transcription in the mammalian transcriptome. *Science* **309**, 1564–1566 (2005).
2. Haddad, F., Bodell, P. W., Qin, A. X., Giger, J. M. & Baldwin, K. M. Role of antisense RNA in coordinating cardiac myosin heavy chain gene switching. *J. Biol. Chem.* **278**, 37132–37138 (2003).
3. Hang, C. T. *et al.* Chromatin regulation by Brg1 underlies heart muscle development and disease. *Nature* **466**, 62–67 (2010).
4. Hung, T. *et al.* Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nature Genet.* **43**, 621–629 (2011).
5. Lin, M. F., Jungreis, I. & Kellis, M. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics* **27**, i275–i282 (2011).
6. Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M. & Weissman, J. S. The ribosome profiling strategy for monitoring translation *in vivo* by deep sequencing of ribosome-protected mRNA fragments. *Nature Protocols* **7**, 1534–1550 (2012).

7. van Rooij, E. *et al.* Control of stress-dependent cardiac growth and gene expression by a microRNA. *Science* **316**, 575–579 (2007).
8. Miyata, S., Minobe, W., Bristow, M. R. & Leinwand, L. A. Myosin heavy chain isoform expression in the failing and nonfailing human heart. *Circ. Res.* **86**, 386–390 (2000).
9. Lompre, A. M. *et al.* Myosin isoenzyme redistribution in chronic heart overload. *Nature* **282**, 105–107 (1979).
10. Schultz, J. J. *et al.* TGF- $\beta$ 1 mediates the hypertrophic cardiomyocyte growth induced by angiotensin II. *J. Clin. Invest.* **109**, 787–796 (2002).
11. Molkenkin, J. D. & Dorn, G. W. II. Cytoplasmic signaling pathways that regulate cardiac hypertrophy. *Annu. Rev. Physiol.* **63**, 391–426 (2001).
12. López, B. *et al.* Osteopontin-mediated myocardial fibrosis in heart failure: a role for lysyl oxidase? *Cardiovasc. Res.* **99**, 111–120 (2013).
13. Frey, N. & Olson, E. N. Cardiac hypertrophy: the good, the bad, and the ugly. *Annu. Rev. Physiol.* **65**, 45–79 (2003).
14. Guttman, M. *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223–227 (2009).
15. Rando, O. J. & Chang, H. Y. Genome-wide views of chromatin structure. *Annu. Rev. Biochem.* **78**, 245–271 (2009).
16. Hahn, M. A., Wu, X., Li, A. X., Hahn, T. & Pfeifer, G. P. Relationship between gene body DNA methylation and intragenic H3K9me3 and H3K36me3 chromatin marks. *PLoS ONE* **6**, e18844 (2011).
17. Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
18. Musselman, C. A. *et al.* Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. *Nature Struct. Mol. Biol.* **19**, 1266–1272 (2012).
19. Liu, R. *et al.* Regulation of CSF1 promoter by the SWI/SNF-like BAF complex. *Cell* **106**, 309–318 (2001).
20. Muchardt, C. & Yaniv, M. A human homologue of *Saccharomyces cerevisiae* SNF2/SWI2 and *Drosophila* brm genes potentiates transcriptional activation by the glucocorticoid receptor. *EMBO J.* **12**, 4279–4290 (1993).
21. Szabó, G. *et al.* Poly(ADP-ribose) polymerase inhibition reduces reperfusion injury after heart transplantation. *Circ. Res.* **90**, 100–106 (2002).
22. Hesselberth, J. R. *et al.* Global mapping of protein–DNA interactions *in vivo* by digital genomic footprinting. *Nature Methods* **6**, 283–289 (2009).
23. Gupta, M. P. Factors controlling cardiac myosin-isoform shift during hypertrophy and heart failure. *J. Mol. Cell. Cardiol.* **43**, 388–403 (2007).
24. Clapier, C. R. & Cairns, B. R. The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* **78**, 273–304 (2009).
25. Jankowsky, E. & Fairman, M. E. RNA helicases—one fold for many functions. *Curr. Opin. Struct. Biol.* **17**, 316–324 (2007).
26. Mallam, A. L., Del Campo, M., Gilman, B., Sidote, D. J. & Lambowitz, A. M. Structural basis for RNA-duplex recognition and unwinding by the DEAD-box helicase Mss116p. *Nature* **490**, 121–125 (2012).
27. Dürr, H., Korner, C., Müller, M., Hickmann, V. & Hopfner, K. P. X-ray structures of the *Sulfolobus solfataricus* SWI2/SNF2 ATPase core and its complex with DNA. *Cell* **121**, 363–373 (2005).
28. Feng, Y. *et al.* Histone H4 acetylation differentially modulates arginine methylation by an *in cis* mechanism. *J. Biol. Chem.* **286**, 20323–20334 (2011).
29. Zuker, M. On finding all suboptimal foldings of an RNA molecule. *Science* **244**, 48–52 (1989).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank C.-H. Chen for assisting with echocardiography; L. Chen, A. Kuo and G. Crabtree for transgene injection and northern blot; M. Ecart and E. Zuo for ribosome analysis. C.-P.C. was supported by the American Heart Association (AHA; Established Investigator Award 12EIA8960018), National Institutes of Health (NIH; HL118087, HL121197), March of Dimes Foundation (#6-FY11–260), California Institute of Regenerative Medicine (CIRM; RN2-00909), Oak Foundation, Baxter Foundation, Stanford Heart Center Research Program, Indiana University (IU) School of Medicine—IU Health Strategic Research Initiative, and the IU Physician-Scientist Initiative, endowed by Lilly Endowment. W.L. and Y.X. were supported by the Oak Foundation; Y.X. by the AHA and Lucile Packard Children's Foundation; C.S. by an NIH fellowship; T.Q. by NIH (HL109512); H.-S.V.C. by CIRM (RB2-01512, RB4-06276) and NIH (HL105194); P.-S.C. by NIH (HL78931, HL71140); B.Z. by NIH (HL116997, HL111770).

**Author Contributions** C.-P.C. and P.H. were responsible for the original concepts, design and manuscript preparation. W.L. and C.-H.L. contributed equally to the work. P.H. conducted most experiments; W.L. and J.Y. assisted with TAC, echo and reporter analyses; C.-H.L. assisted with protein purification; S.T.N. assisted with ribosome data analysis; K.K.J. assisted with protein sequence and motif analysis; C.S. assisted with western blot studies; W.X. assisted with CSF scoring; Y.X. assisted with RNA/protein staining; C.-J.L. and C.-Y.L. assisted with *Brg1*-null tissue preparation and H-10 antibody-ChIP optimization; H.-C.C. assisted with cloning; H.-S.V.C. assisted with tissue collection; E.A. assisted with tissue collection/rat tissue supply; B.Z. assisted with driver line generation; D.B., P.-S.C. and T.Q. assisted with data analysis.

**Author Information** Data have been deposited in the Gene Expression Omnibus under accession number GSE49716. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.-P.C. ([changcp@iu.edu](mailto:changcp@iu.edu)).



## METHODS

**Mice, animal sample size, and randomization.** For the generation of *Tg779* mice, *Mhrt779* was cloned into the pTRE2 backbone (Clontech). A DNA fragment containing the Tre promoter and *Mhrt779* were injected into the pronucleus of fertilized oocytes (B6C3H/F1). Embryos were implanted into a pseudopregnant CD-1 mouse. The *Tre-Mhrt779* transgene was identified by PCR genotyping using primers CGCCTGGAGACGCCATCCAC and TGTCTTCAAAGCTGACTCCCT. *Tre-Mhrt779* mice with ~3 copies of the transgene were backcrossed with *Tnnt2-rtTA* mice as described previously<sup>3,30</sup> to generate *Tnnt2-rtTA;Tre-Mhrt779* (*Tg779*) mice. The number of animals used (*n*) is denoted in each test in the figures, including technical replicates when applicable. We routinely used mouse littermates to control and perform our experiments. Each subgroup of experiments had *n* = 3 to 14 biological replicates, many of which had technical replicates of three. Assignment to each experimental subgroup was based on genotypes. Littermate mice with the same genotypes regardless of gender were randomly selected from the cage and assigned to different control and experimental subgroups. Major procedures were blinded. The use of mice for studies was in compliance with the regulations of Indiana University, Stanford University and the National Institutes of Health.

**RACE and cloning of full length of *Mhrt* transcripts.** The 3' and 5' RACE were performed using the FirstChoice RLM-RACE Kit (Ambion) following the manufacturer's instruction. RNA was extracted from adult heart ventricles. Primers used for 3' and 5' RACE were designed based on the known sequence information: TCATTGGCAGGACAGCATC (first-round *Mhrt* 3'-prime specific) and GAGCA TTTGGGATGGTATAC (second-round *Mhrt* 3'-prime specific); CAACACTT TTCATTTCTCTTT (first-round *Mhrt* 5'-prime specific) and TCTGCTTCA TTGCTCTGTTT (second-round *Mhrt* 5'-prime specific). Once we reached the 5' and 3' cDNA ends, we used primers F1 (Fig. 1a; AAGAGCCCTACAGTCTG ATGAACA) and R1 (Fig. 1a; CCTTCACACAAACATTTTATTT) to amplify the full-length *Mhrt* transcripts and cloned them into pDrive TA cloning vector (Qiagen) for sequencing. *Mhrt* RNAs were also further cloned into shuttle vector pAdd2 (refs 31, 32) for expression in cells.

**Northern blot and *in situ* hybridization.** We obtained 5 µg of total RNA using Quick-RNA Mini Kit (Zymo Research). RNA blot was performed using Northern-Max Kit (Ambion) following the manufacturer's protocol. Single-stranded RNA probe was generated by *in vitro* transcription with Maxlscript SP6/T7 kit (Ambion) with ATP [ $\alpha$ -<sup>32</sup>P] (PerkinElmer) using full-length *Mhrt779*, *Myh6* and *Myh7* as the template and followed by digestion with DNase I (Ambion). Hybridization was performed at 65 °C. The blot was washed and imaged by Phosphor storage scanning by Typhoon 8600 Imager (GE Healthcare). *In situ* hybridization experiments were performed as previously described<sup>3,33</sup>.

**RNA fractionation.** To isolate cytosolic and nuclear RNAs from adult heart tissues, we used a PARIS kit (Ambion) and followed the manufacturer's instruction. Ten milligrams of tissue were homogenized in cell fractionation buffer thoroughly before centrifuging for 5 min at 500g. Supernatant was collected as the cytosolic fraction, while the nuclear pellet was washed and lysed by cell disruption buffer. Such samples were further mixed with 2× lysis/binding solution before extracting RNA using the manufacturer's protocol.

**Codon substitution frequency predication.** To measure the coding potential of *Mhrt*, we used the previously described codon substitution frequencies (CSF) method<sup>4,5</sup> to evaluate the evolutionary characteristics in their alignments with orthologous regions in six other sequenced mammalian genomes (rat, human, hamster, rhesus monkey, cat and dog). CSF generates a likelihood score for a given sequence considering all codon substitutions observed within its alignment across multiple species, which was based on the relative frequency of similar substitutions occurring in known coding and noncoding regions. CSF compares two empirical codon models; one generated from alignments of known coding regions and the other according to noncoding regions, producing a likelihood ratio. The ratio reflects whether the protein-coding model better explains the alignment.

**Ribosome profiling and RNA deep sequencing.** For ribosome profiling<sup>6</sup>, overexpression of the predominant species of *Mhrt* (*Mhrt779*) along with *HOTAIR* were achieved through co-transfecting pAdd2-779 and pAdd2-HOTAIR into SW13 cells. The cells were then lysed to extract ribosome-associated RNA fragments using ARTseq Ribosome Profiling Kit (Epicentre, Illumina). The RNA fragments were further converted into a DNA library through end repair, adaptor ligation, reverse transcription circularization, and PCR amplification. A conventional RNA-seq library was also prepared, with total RNA extracted from those cells with an miRNeasy Mini Kit (Qiagen #217004). The libraries were further processed according to a MiSeq Sample Prep sheet, and a MiSeq 50 cycle kit was used for sequencing. PCR products (1.25 pmol) were used for sequencing. Approximately 600,000–700,000 reads were properly paired and used for further analysis. The resulting reads were aligned to the human hg19 or mouse mm10 genome using Bowtie2 v.2.0.0.6 (ref. 34). Mapped reads were visualized on the UCSC browser as bigwig files generated using samtools v.0.1.18 (ref. 35), bedtools v.2.16.1 (ref. 36), bedClip and bedGraphToBigWig.

For quantification of fragments per kilobase of exon per million fragments mapped (FPKM) values, cuffdiff as part of the tophat suite v.2.0.8b<sup>37</sup> was run on a merged bam file containing the human and the *Mhrt* reads using a custom gtf file comprising the human hg19 iGenome and the *Mhrt* transcripts. To generate scatter plots of the genes, cuffdiff files were used for visualization with cummerbund v.2.3.1 (ref. 37).

***In vitro* translation and biotin labelling.** TNT Quick Coupled Transcription/Translation System (Promega) was used for *in vitro* translation. Briefly, 1 µg plasmids of control (luciferase) and various *Mhrt* species inserted into pDrive vector were added to 40 µl rabbit reticulocyte lysates containing <sup>35</sup>S-methionine. After 1 h of incubation, the reactions were analysed on 10–20% Tris-Tricine gel. The gel was dried and visualized by the Typhoon 8600 Imager (GE Healthcare). Biotin-NTP was added to the *in vitro* translation reaction. Total RNAs were extracted and the biotin-labelled RNAs were detected subsequently by IRDye 680 Streptavidin (LICOR, 926-68079) using an Odyssey Infrared Imaging System.

**TAC.** The TAC surgery was performed as described<sup>3</sup> on adult mice of 8–10 weeks of age and between 20 and 25 g in weight. Mice were fed with doxycycline food pellets (6 mg doxycycline per kg of food; Bioserv) 7–14 days before the TAC operation. Mice were anaesthetized with isoflurane (2–3%, inhalation) in an induction chamber and then intubated with a 20-gauge intravenous catheter and ventilated with a mouse ventilator (Minivent, Harvard Apparatus). Anaesthesia was maintained with inhaled isoflurane (1–2%). A longitudinal 5 mm incision of the skin was made with scissors at the midline of sternum. The chest cavity was opened by a small incision at the level of the second intercostal space 2–3 mm from the left sternal border. While opening the chest wall, the chest retractor was gently inserted to spread the wound 4–5 mm in width. The transverse portion of the aorta was bluntly dissected with a curved forceps. Then, 6-0 silk was brought underneath the transverse aorta between the left common carotid artery and the brachiocephalic trunk. One 27-gauge needle was placed directly above and parallel to the aorta. The loop was then tied around the aorta and needle, and secured with a second knot. The needle was immediately removed to create a lumen with a fixed stenotic diameter. The chest cavity was closed by 6-0 silk suture. Sham-operated mice underwent similar surgical procedures, including isolation of the aorta and looping of the aorta, but without tying of the suture. The pressure load caused by TAC was verified by the pressure gradient across the aortic constriction measured by echocardiography. Only mice with a pressure gradient >30 mm Hg were analysed for cardiac hypertrophy, echocardiography and other purposes.

**Echocardiography.** The echocardiographer was blinded to the genotypes and surgical procedure. Transthoracic ultrasonography was performed with a GE Vivid 7 ultrasound platform (GE Health Care) and a 13 MHz transducer was used to measure aortic pressure gradient and left ventricular function. Echocardiography was performed on control and *Tnnt2-rtTA;Tre-Mhrt779* (*Tg779*) mice at designated time points after the TAC procedure. To minimize the confounding influence of different heart rates on the aortic pressure gradient and left ventricular function, the flow of isoflurane (inhalational) was adjusted to anaesthetize the mice while maintaining their heart rates at 450–550 beats per minute. The peak aortic pressure gradient was measured by continuous-wave Doppler across the aortic constriction. Left ventricular function was assessed by M-mode scanning of the left ventricular chamber, standardized by two-dimensional, short-axis views of the left ventricle at the mid papillary muscle level. Left ventricular chamber size and wall thickness were measured in at least three beats from each projection and averaged. Left ventricular internal dimensions at diastole and systole (LVIDd and LVIDs, respectively) were measured. The fractional shortening (FS) of the left ventricle was defined as  $100\% \times (1 - \text{LVIDs}/\text{LVIDd})$ , representing the relative change of left ventricular diameters during the cardiac cycle. The mean FS of the left ventricle was determined by the average of FS measurements of the left ventricular contraction over five beats. *P* values were calculated by Student's *t*-test. Error bars indicate s.e.m.

**Histology, trichrome staining and morphometric analysis of cardiomyocytes.** Histology and trichrome staining were performed as described<sup>38,39</sup>. Trichrome stain (Masson) kit (Sigma) was used and the manufacturer's protocol was followed. For morphometric analysis of cardiomyocytes, paraffin sections of the heart were immunostained with a fluorescein isothiocyanate-conjugated wheat germ agglutinin (WGA) antibody (F49, Biomedex) that highlighted the cell membrane of cardiomyocytes. Cellular areas outlined by WGA were determined by the number of pixels enclosed using ImageJ software (NCBI). Approximately 250 cardiomyocytes of the papillary muscle at the mid-left ventricular cavity were measured to determine the size distribution. *P* values were calculated by Student's *t*-test. Error bars indicate s.e.m.

**RT-qPCR and strand-specific reverse transcription PCR analysis.** RT-qPCR analyses were performed as described<sup>3,38</sup>. The following primer sequences (listed later) were used. RT-qPCR reactions were performed using SYBR green master mix (BioRad) with an Eppendorf realplex, and the primer sets were tested to be quantitative. Threshold cycles and melting curve measurements were performed with software. *P* values were calculated by Student's *t*-test. Error bars indicate s.e.m. To conduct strand-specific RT-PCR analysis, human total RNA and Superscript III First-Strand



Synthesis System (Invitrogen) was used. Primers R1 (Fig. 4k; CTACAGAATGAG ATCGAGGACT) and R2 (Fig. 4k; GGGGCTGAAGAGTGAGCCTT) were designed based on known sequence and were used for individual RTs, respectively. To detect *MHRT*, primers F1 (Fig. 4k; CTGGAGCTGGGACAGGTCAGCA) and R1 were used. These primers could also amplify endogenous *MYH7* and thus serve as controls. Primers F2 (Fig. 4k; TGGGGAACACGGCGTTCTTGA) and R2 were used to specifically amplify *MHRT* and used in RT-qPCR analysis.

PCR primers for RT-qPCR of mRNA were as follows. Mouse *Tf11b*-F, CTCTG TGGCGGCAGCAGCTATTT, mouse *Tf11b*-R, CGAGGGTAGATCAGTCTGTA GGA; mouse *Hprt1*-F, GCTGGTGAAGAGGACCTCT, mouse *Hprt1*-R, CACAG GACTAGAACACCTGC; mouse *Anf*-F, GACTAGGCTGCAACAGCTTCCG, mouse *Anf*-R, GCCACAGTGGCAATGTGACCAA; mouse *Serca2a*-F, CATTTC CATTGACGTCTGGAT, mouse *Serca2a*-R, CTTTGCCATCTACGAGTTCC; mouse *Tnnt2*-F, TACAGACTCTGATCGAGGCTCACTTC, mouse *Tnnt2*-R, TC ATTGCGAATACGCTGCTGCTC; mouse *Mhrt*-F (common), GAGCATTTGG GGATGGTATAC, mouse *Mhrt*-R (common), TCTGCTTCATTGCCTCTGTT T; mouse *Mhrt779*-F, TCTGGCCACAGCCCGCAGCTTC, mouse *Mhrt779*-R, AGTCATGTATACCATCCCCAA; mouse *Neat1*-F, TCTCCTGGAGCCACATC TCT, mouse *Neat1*-R, GCTTTTCTTAGGCCCAAAC; mouse 28S-rRNA-F, GG TAGCCAAATGCCTCGTCAT, mouse 28S-rRNA-R, CCCTTGGCTGTGGTTT CG; human *TFIIB*-F, ACCACCCCAATGGATGCAGACAG, human *TFIIB*-R, A CGGGCTAAGCGTCTGGAC; human *MHRT*-F (F2), TGGGGAACACGGCG TTCTTGA, human *MHRT*-R (R2), GGGGCTGAAGAGTGAGCCTT; human *HOTAIR*-F, GGTAGAAAAAGCAACCAAGCAAGC, human *HOTAIR*-R, ACAT AAACCTCTGTCTGTGAGTGCC; human *GAPDH*-F, CCGGGAAACTGTGG CGTGATGG, human *GAPDH*-R, AGGTGGAGGAGTGGGTGTCGCTGTT.

**ChIP-qPCR.** ChIP assay was performed as described<sup>3</sup> with modifications. Briefly, chromatin from hearts or SW13 cells was sonicated to generate average fragment sizes of 200–600 bp, and immunoprecipitated using anti-BRG1 J1 antibody<sup>3,40</sup>, anti-Brg1 H-10 antibody (Santa Cruz Biotechnology, against 115–149 amino acids of N terminus Brg1), anti-RNA polymerase II (Pol II) antibody (ab24759, Abcam), anti-H3K4me3 antibody (07-473, Millipore), anti-H3K36me3 antibody (17-10032, Millipore) or normal control IgG. Isolation and purification of immunoprecipitated and input DNA were done according to the manufacturer's protocol (Magna ChIP Protein G Magnetic Beads, Millipore), and qPCR analysis of immunoprecipitated DNA were performed. ChIP-qPCR signal of individual ChIP reactions was standardized to its own input qPCR signal or IgG ChIP signal. PCR primers (listed later) were designed to amplify the promoter regions of mouse *Myh6* (–426, –320), mouse *Myh7* (–102, +58), mouse *Shh* (–7142, –6911), mouse *Vegfa* (+1, +150) human *GAPDH* (–45, +121). The DNA positions are denoted relative to the transcriptional start site (+1).

PCR primers for ChIP-qPCR are as follows. Mouse ChIP-*Myh6* promoter-F, GCAGATAGCCAGGGTTGAAA, mouse ChIP-*Myh6* promoter-R, TGGGTAA GGGTCACCTTCTC; mouse ChIP-*Myh7* promoter-F, GTGACACAGCCCT TTCTAAAT, mouse ChIP-*Myh7* promoter-R, CTCCAGCTCCCACTCCTACC; mouse ChIP-*Shh* promoter-F, GAGAACATTACAGGGTAGGAA, mouse ChIP-*Shh* promoter-R, GAAGCAGTGAGGTGGTGG; mouse ChIP-*Vegfa* promoter-F, CAAATCCAGAGCACAGACTC, mouse ChIP-*Vegfa* promoter-R, AGCGCAG AGGCTTGGGGCAGC; human ChIP-*GAPDH* promoter-F, TACTAGCGGTTTT ACGGGCG, human ChIP-*GAPDH* promoter-R, TCGAACAGGAGGAGCAGAG AGCGA.

**RNA immunoprecipitation.** RNA immunoprecipitation (RNA-IP, RIP) was conducted as described<sup>4</sup> with some modifications. Briefly, P1 hearts, sham hearts or those from mice that had undergone TAC, or SW13 cells were crosslinked and lysed with lysis buffer (10 mM HEPES pH 7.5, 85 mM KCl, 0.5% NP-40, 1 mM dithiothreitol (DTT), 1× protease inhibitor) for tissues or lysis buffer (10 mM Tris-HCl pH 8.1, 10 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 0.5% NP-40, 1 mM DTT, 1× protease inhibitor) for cells. Nuclei were isolated and sonicated using Bioruptor (Diagenode) (30 s on, 30 s off, power setting H, 5 min, performed twice) in nuclear lysis buffer (50 mM Tris-HCl pH 8.1, 150 mM NaCl, 0.1% NP-40, 1 mM DTT, protease inhibitor, ribonuclease inhibitor). The nuclear extract was collected and incubated with primary antibodies at 4 °C overnight together with Manga ChIP Protein G Magnetic Beads (Millipore). The beads were washed by wash buffer I (20 mM Tris-HCl pH 8.1, 150 mM NaCl, 1% Triton X-100 and 0.1% SDS) three times, and wash buffer II (20 mM Tris-HCl pH 8.1, 500 mM NaCl, 1% Triton X-100 and 0.1% SDS) three times. Beads were then resuspended in 150 µl 150 mM RIPA (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.1% SDS, 1% NP-40, 0.5% sodium deoxycholate) with 5 µl Proteinase K and incubated for 1 h at 65 °C. We added 1 ml of TRIzol to the sample, and RNA was extracted using the Quick-RNA Mini Kit with the on-column DNase I digest (ZymoResearch). RT and qPCR were then conducted with the purified RNA. The antibodies used for the immunoprecipitation are anti-BRG1 J1 antibody<sup>3,40</sup>, Ezh2 (ref. 41) (Active Motif), Suz12 (refs 41, 42) (Bethyl Laboratories) and normal IgG control.

**Reporter assay and truncation of the *Mhrt* promoter.** For the *Mhrt* promoter reporter assay, plasmid was constructed by inserting ~2.5 kb mouse *Mhrt* promoter into the episomal pREP4-Luc plasmid<sup>3,19,38,43</sup> through cloning the PCR-amplified region of the promoter by using primers ACCGGCTGAACCCCACT TCC and ATGTCGAGACAGGGAACAGAA. Mouse *Myh6* (–426 to +170, based on new genome annotation) and *Myh7* (–3561 to +222) reporter constructs were described previously<sup>3</sup>. These vectors were transfected into rat neonatal cardiomyocytes or SW13 cells using lipofectamine 2000 (Invitrogen) along with plasmids expressing mouse Brg1 (actin-mBrg1-IRES-eGFP) or a matching empty vector plasmid (gifts from G. Crabtree) as well as an episomal *Renilla* luciferase plasmid (pREP7-RL) to normalize transfection efficiency. The transfected cells were cultured for 48 h and harvested for luciferase assay using the dual luciferase assay kit (Promega). For naked DNA reporter, mouse *Myh6* promoter (–426 to +170) was inserted in pGL3 vector (Promega), and *Renilla* luciferase plasmid pRL-SV40 (gifts from J. Chen) was used as a normalizer. Dual luciferase assay was performed according to the manufacturer's instruction 48 h after transfection. For deletion analysis of the *Mhrt* promoter, various regions of the promoter were deleted from the full-length pREP4-*Mhrt*. The constructs were further analysed by transfecting into SW13 cells.

**RNA-EMSA and  $K_d$  calculation.** Biotin-labelled RNA probe was generated by *in vitro* transcription with MAXIscript SP6/T7 kit (Ambion) with biotin labelling NTP mixture (Roche) using linearized pDrive-*Mhrt779* construct as the template and followed by digestion with DNase I (Ambion). EMSA was performed by using the LightShift Chemiluminescent RNA EMSA Kit (Thermo Scientific). The labelled probe was incubated with appropriate amounts of recombinant proteins in 10 µl in the 1× binding buffer (10 mM HEPES-KOH, pH 7.3, 10 mM NaCl, 1 mM MgCl<sub>2</sub>, 1 mM DTT) with 5 µg tRNA carrier at room temperature for 30 min. The reactions were then loaded onto 1% 0.5× TBE agarose gel and transferred to BrightStar-Plus positive charged membrane. The biotin-labelled probes were detected and quantified subsequently by IRDye 680 Streptavidin (Li-COR, 926-32231) using Odyssey Infrared Imaging System. The shifted signals were quantified and plotted against amount of the MBP, MBP-D1, MBP-D2 and MBP-D1D2 proteins using a previously described method<sup>26</sup> with GraphPad Prism (GraphPad). The software facilitates the fitting of nonlinear regression model and calculation of  $K_d$  values based on the fitting curve. The errors and  $r^2$  values were also generated from the fitting curve.

**Protein expression and purification of Brg1 helicase domains.** To generate MBP fusion proteins of mouse Brg1 helicase domains, the DExx-box domain (D1) (amino acids 774–913 of Brg1), helicase-C domain (D2) together with C-terminal extension (CTE) (amino acids 1086–1310 of Brg1), as well as the entire helicase region (D1D2) (774–1310) were amplified by PCR and cloned into pMAL vector. MBP fusion proteins were induced by isopropyl-β-D-thiogalactoside (IPTG) and purified by amylose resin (E8021S, NEB).

**Nucleosome assembly and amylose pull-down.** Nucleosome assembly was performed by using EpiMark Nucleosome Assembly Kit (E5350S, NEB) following the manufacturer's instruction<sup>28</sup>. In brief, recombinant human core histone octamer, which consists of the 2:1 mix of histone H2A/H2B dimer and histone H3.1/H4 tetramer, were mixed with purified 5S rDNA (208 bp; N1202S, NEB), *Neo* (512 bp, amplified from pST18-*Neo*; 117502S, Roche), *Myh6* core promoter (596 bp, –426 to +170) and *Mhrt* core promoter (a3/a4, 596 bp, –2290 to –1775) DNA at 2 M NaCl. PCR primers to amplify *Neo* are CGATCGCTGCGAATCGGGA and CACTGA AGCGGGAAGGGACT. The salt concentration was gradually lowered by dilution to allow the formation of nucleosomes. The EMSA assay was used to assess the efficiency of nucleosome assembly. For amylose pull-down assay, the amylose resin (E8021S, NEB) was washed thoroughly and equilibrated with binding buffer (10 mM Tris-HCl, pH 7.5, 150 mM NaCl) before incubation with purified MBP or MBP-D1D2 proteins for 2 h. Nucleosome mixture or naked DNA mixture of 5S rDNA, *Neo* and *Myh6* promoter DNA were added for incubation at 4 °C for overnight. The resin was then washed excessively by washing buffer (20 mM Tris-HCl, pH 8.1, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS) before decross-linking and extraction of the DNA with phenol:chloroform:isoamyl alcohol. For competition assays, *in vitro* transcribed *Mhrt779* was incubated with MBP-D1D2 in binding buffer (10 mM HEPES-KOH, pH 7.3, 10 mM NaCl, 1 mM MgCl<sub>2</sub>, 1 mM DTT) with ribonuclease inhibitor at room temperature for 30 min before adding nucleosomal DNA. The subsequent incubation, wash and DNA purification were performed as regular amylose pull-down assays. The qPCR signal of individual pull-down reaction was standardized to its own input RT-qPCR signal. qPCR primers were designed to amplify the 5S rDNA (CAAGCAAGAGCCTACGACCA; ATTC GTTGAATTCCTCGGG), *Neo* (TAAAGCACGAGGAAGCGGTC; TCGACCC CAAGCGAAACAT), *Myh6* promoter (GCAGATAGCCAGGGTTGAAA; TGGG TAAGGGTCACCTTCTC) and *Mhrt* promoter (ATGCCAAATGGTTGCTCTTT; GAGCTTGAGAACAGGCAGT).

**Cloning of Brg1 truncation constructs.** For cloning of truncated Brg1 with deletion of amino acids 774–913 ( $\Delta$ D1) or 1086–1246 ( $\Delta$ D2), primers with an NheI restriction digestion site, which complement the downstream and upstream sequences of the truncated region ( $\Delta$ D1: CCCGGGGCTAGCCTGCAGAACAGCTACCGGAGCT and CCCGGGGCTAGCCAGGTTGTTGTGTACAGG GACA;  $\Delta$ D2: CCCGGGGCTAGCATCAAGAAGTTCAAATTTCCC and CCCGGGGCTAGCCTGCAGGCCATCCTGGAGCAGCAGCAG) were used to amplify from pActin-Brg1-IRES-eGFP by KOD Xtreme Hot Start DNA Polymerase (Novagen). After digestion with NheI, the linearized fragment was subject to ligation and transformation. The truncation constructs were sequenced to confirm the fidelity of the cloning. Western blot was further performed to assess the expression of the constructs. Monoclonal H-10 antibodies (Santa Cruz Biotech, sc-374197), which were raised against Brg1 N-terminal amino acids, were used in the experiments involving truncated Brg1.

**Protein sequence analysis.** Brg1 core helicase domain (774–1202) was applied for secondary structure prediction using the Fold & Function Assignment System (FFAS) server (<http://ffas.burnham.org/ffas.cgi/cgi/ffas.pl>). The output revealed that Brg1 core helicase domains are structural homologues of SF2 helicases: Vasa<sup>44</sup> (fruit fly, Protein Data Bank (PDB) accession number 2DB3), Rad54 (refs 27, 45) (zebrafish PDB accession 1Z3I, *Sulfolobus solfataricus* PDB accession 1Z63) and Chd1 (ref. 46) (yeast, PDB accession 3MWY). Those proteins, together with Brg1, were further employed for multiple sequence alignment with T-Coffee, which is a program allowing combination of the results obtained with several alignment methods (<http://www.ebi.ac.uk/Tools/msa/tcoffee/>).

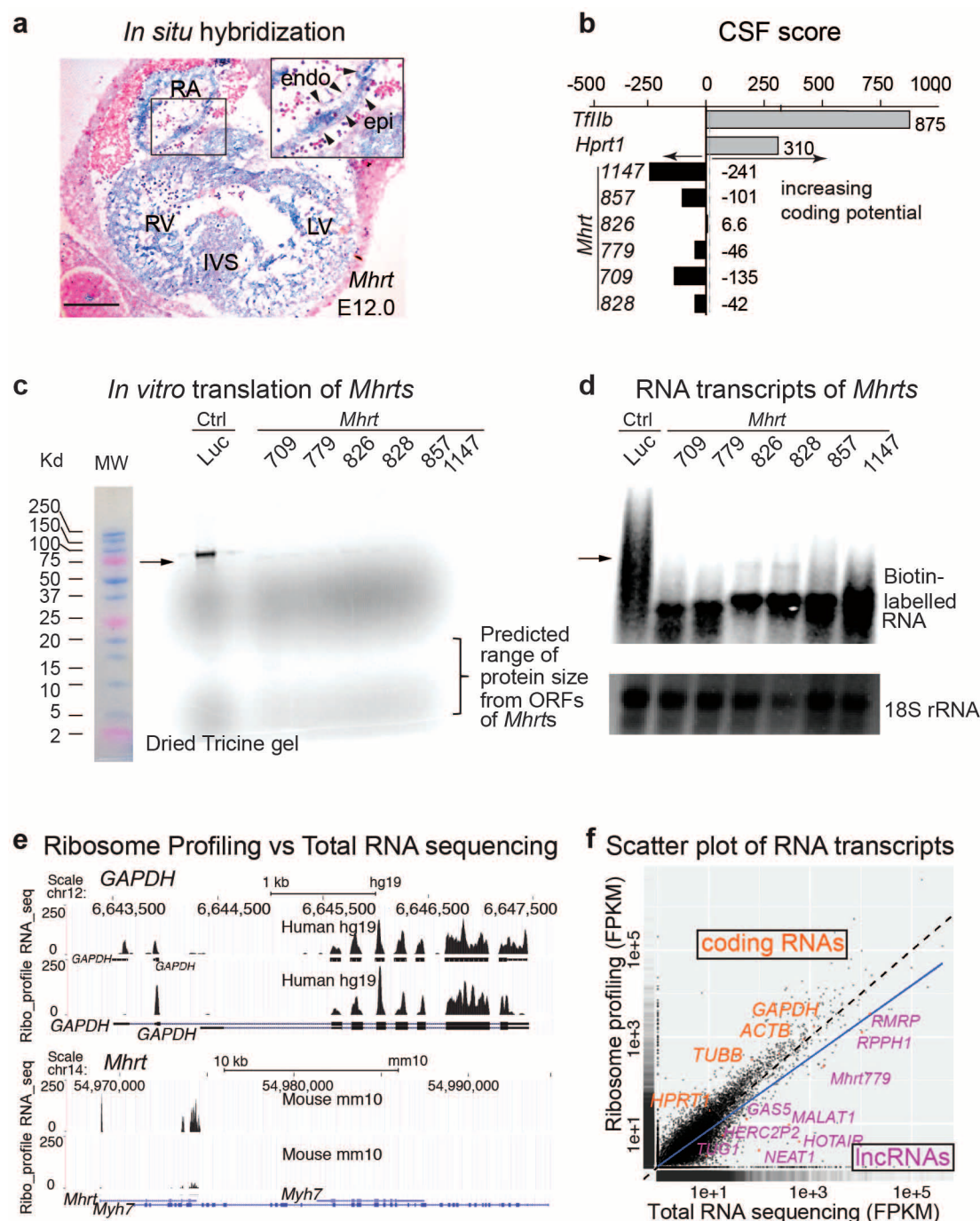
**RNA secondary structural prediction.** To predict the secondary structure for mouse *Mhrt* and human *MHRT*, the single-stranded sequence of *Mhrt*<sup>779</sup> and human *MHRT* were analysed on the Vienna RNAfold web server (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>) with calculation of minimum free energy<sup>29,47–49</sup>.

**Human heart tissue analysis.** Human tissues were processed for RT–qPCR and strand-specific RT–PCR. The use of human tissues is in compliance with the regulation of Sanford/Burnham Medical Research Institute, Intermountain Medical Center, Stanford University, and Indiana University.

**Primary cardiomyocyte culture.** For functional studies in cardiomyocytes, neonatal rat ventricular cardiomyocytes were cultured as previously described<sup>150,51</sup>. Briefly, P0 or P1 Sprague–Dawley rats were used. The ventricles were excised and trypsinized for 15 min 4–5 times. Cells were then collected and resuspended in DMEM supplements with 10% FBS. The cells were plated for 1 h to allow the attachment of non-cardiomyocyte cells. The remaining cardiomyocytes were plated at a density of  $2 \times 10^5$  cells ml<sup>-1</sup>. The cells were transfected with Lipofectamine 2000 (Invitrogen) after 48 h.

30. Wu, B. *et al.* Inducible cardiomyocyte-specific gene disruption directed by the rat Tnnt2 promoter in the mouse. *Genesis* **48**, 63–72 (2010).

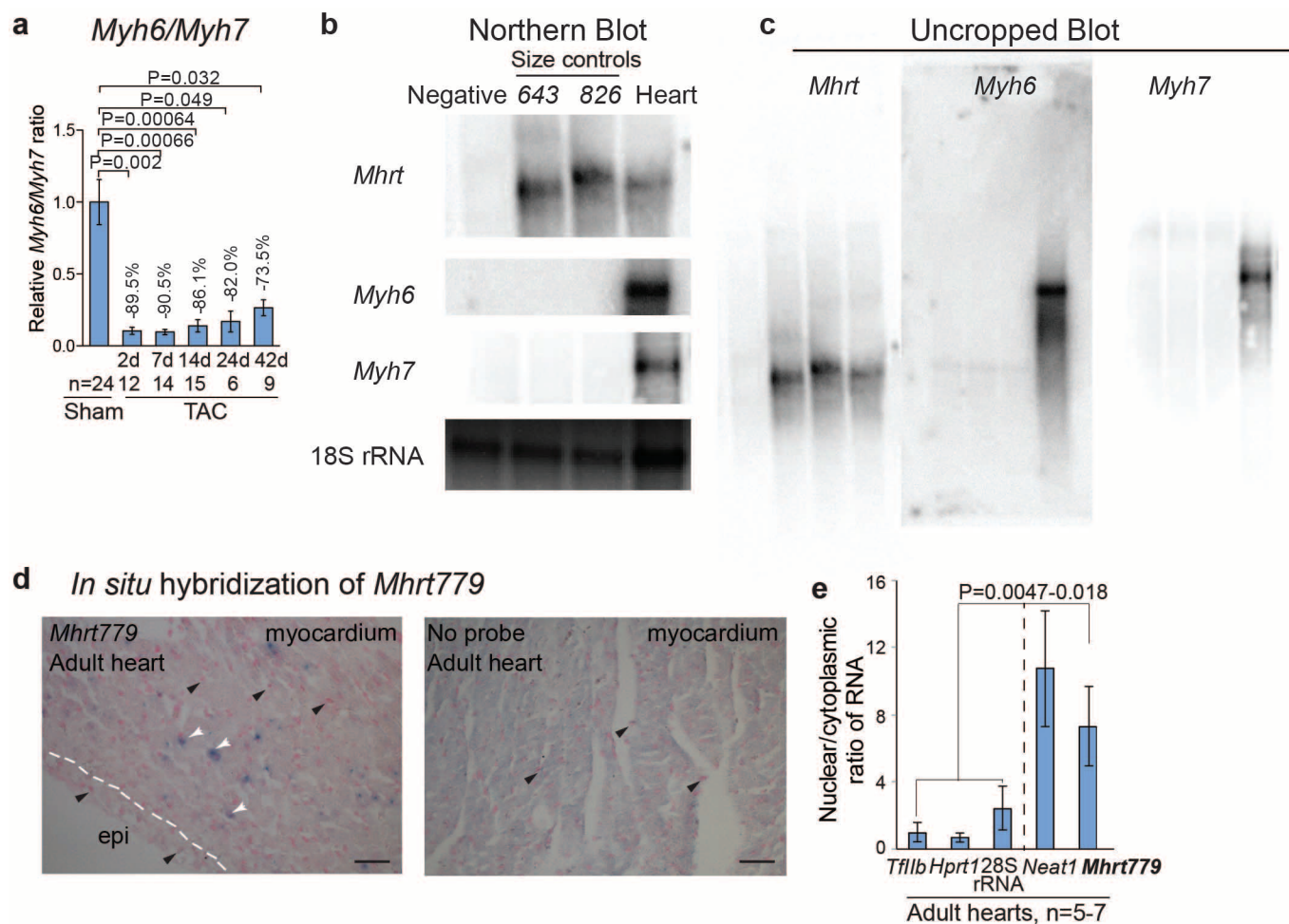
31. Wei, K., Kuhnert, F. & Kuo, C. J. Recombinant adenovirus as a methodology for exploration of physiologic functions of growth factor pathways. *J. Mol. Med. (Berl.)* **86**, 161–169 (2008).
32. Kuhnert, F. *et al.* Essential regulation of CNS angiogenesis by the orphan G protein-coupled receptor GPR124. *Science* **330**, 985–989 (2010).
33. Xiong, Y. *et al.* Brg1 governs a positive feedback circuit in the hair follicle for tissue regeneration and repair. *Dev. Cell* **25**, 169–181 (2013).
34. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
35. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
36. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
37. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7**, 562–578 (2012).
38. Stankunas, K. *et al.* Endocardial Brg1 represses *ADAMTS1* to maintain the microenvironment for myocardial morphogenesis. *Dev. Cell* **14**, 298–311 (2008).
39. Chang, C. P. *et al.* A field of myocardial-endocardial NFAT signaling underlies heart valve morphogenesis. *Cell* **118**, 649–663 (2004).
40. Khavari, P. A., Peterson, C. L., Tamkun, J. W., Mendel, D. B. & Crabtree, G. R. Brg1 contains a conserved domain of the SWI2/SNF2 family necessary for normal mitotic growth and transcription. *Nature* **366**, 170–174 (1993).
41. Grote, P. *et al.* The tissue-specific lncRNA *Fendrr* is an essential regulator of heart and body wall development in the mouse. *Dev. Cell* **24**, 206–214 (2013).
42. Klattenhoff, C. A. *et al.* *Braveheart*, a long noncoding RNA required for cardiovascular lineage commitment. *Cell* **152**, 570–583 (2013).
43. van der Vlag, J., den Blaauwen, J. L., Sewalt, R. G., van Driel, R. & Otte, A. P. Transcriptional repression mediated by polycomb group proteins and other chromatin-associated repressors is selectively blocked by insulators. *J. Biol. Chem.* **275**, 697–704 (2000).
44. Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S. & Yokoyama, S. Structural basis for RNA unwinding by the DEAD-box protein *Drosophila* Vasa. *Cell* **125**, 287–300 (2006).
45. Thomä, N. H. *et al.* Structure of the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54. *Nature Struct. Mol. Biol.* **12**, 350–356 (2005).
46. Hauk, G., McKnight, J. N., Nodelman, I. M. & Bowman, G. D. The chromodomains of the Chd1 chromatin remodeler regulate DNA access to the ATPase motor. *Mol. Cell* **39**, 711–723 (2010).
47. Zuker, M. & Stiegler, P. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res.* **9**, 133–148 (1981).
48. Gruber, A. R., Lorenz, R., Bernhart, S. H., Neubock, R. & Hofacker, I. L. The Vienna RNA websuite. *Nucleic Acids Res.* **36**, W70–W74 (2008).
49. Wan, Y., Kertesz, M., Spitale, R. C., Segal, E. & Chang, H. Y. Understanding the transcriptome through RNA structure. *Nature Rev. Genet.* **12**, 641–655 (2011).
50. Fu, X. M., Yao, Y. J., Yang, Z., Xiang, L. & Gao, J. [Alteration and its significance to expression of aquaporin-4 in cultured neonatal rat astrocytes in the model of hypoxic damage.] *Sichuan Da Xue Xue Bao Yi Xue Ban* **36**, 641–644 (2005).
51. Yang, J. *et al.* C-reactive protein augments hypoxia-induced apoptosis through mitochondrion-dependent pathway in cardiac myocytes. *Mol. Cell. Biochem.* **310**, 215–226 (2008).



**Extended Data Figure 1 | *Mhrt* RNAs have no coding potential.** **a**, RNA *in situ* analysis of *Mhrt* (blue) in a mouse E12 heart. The RNA probe targets all *Mhrt* species. Red: nuclear fast red. Black arrowheads indicate nuclei of endothelial, endocardial or epicardial cells. Inset shows magnified region from the boxed area. endo, endocardium; epi, epicardium; IVS, interventricular septum; LV, left ventricle; RA and RV, right atrium and ventricle, respectively. Scale bars = 100  $\mu$ m. **b**, Codon substitution frequency (CSF) scores of *Tfib* and *Hprt1* mRNA, as well as full-length *Mhrt* species. **c**, *In vitro* translation of control *Mhrt* species (709, 779, 826, 828, 857, 1147) and luciferase (Luc). Arrow points to the protein product of luciferase. **d**, Biotin-labelling of *Mhrt* species

(709, 779, 826, 828, 857, 1147) and luciferase in the *in vitro* translation reactions. Arrow points to the RNA product of luciferase. **e**, Ribosome profiling relative to whole transcriptome RNA sequencing. x-axis: genomic position at the human *GAPDH* and the murine *Myh7* loci. y-axis: mapped reads. **f**, Scatter plot of RNA in fragments per kilobase per million reads (FPKM). Noncoding RNAs (purple) cluster towards the x-axis; coding RNAs (orange) towards the y-axis. *Mhrt779* is found below both the identity line (dashed, slope = 1, intercept = 0) and the smooth-fit regression line (in blue). RNA examples are endogenous except that *HOTAIR* was co-transfected with *Mhrt779*.

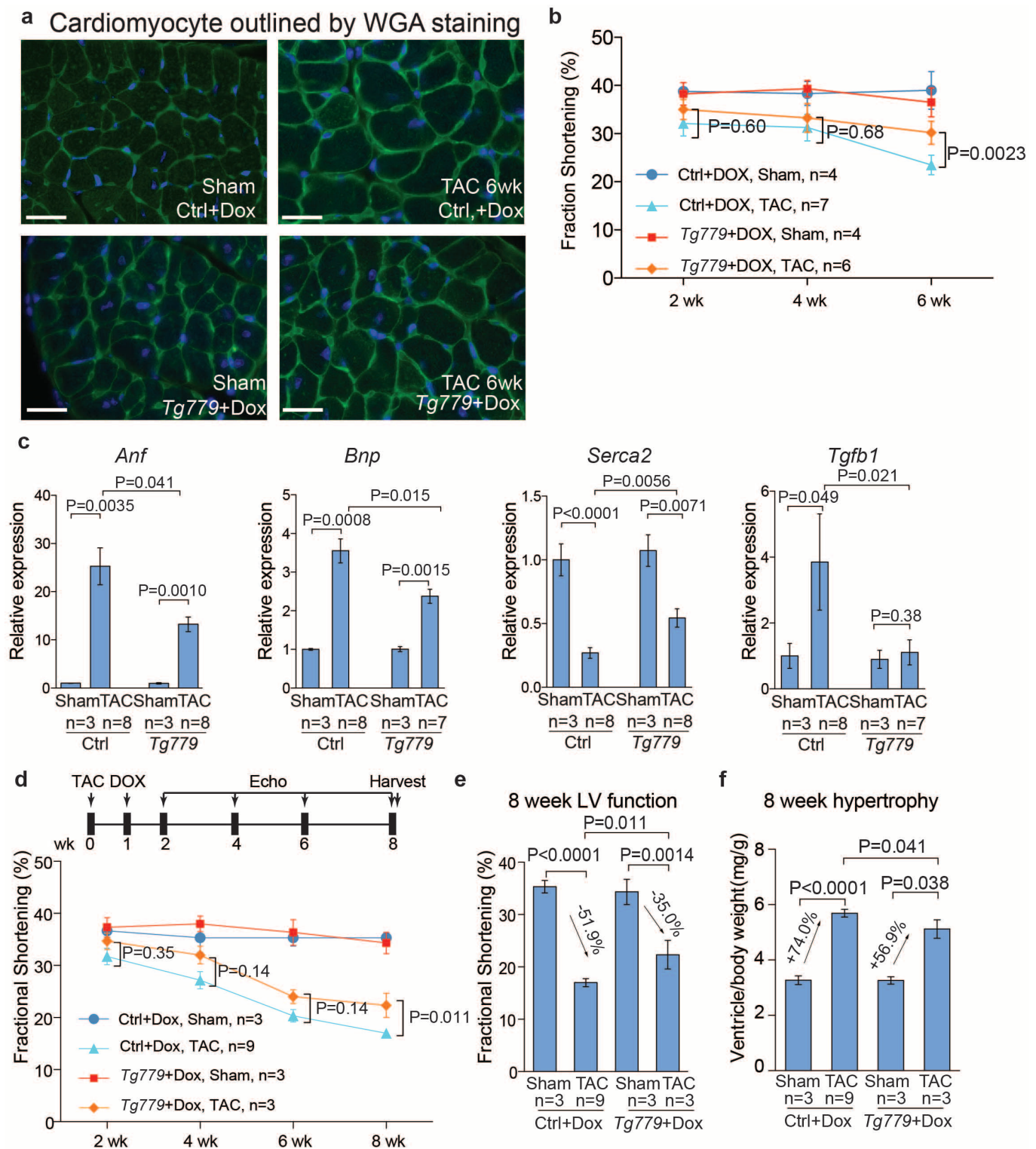




**Extended Data Figure 2 | Quantification of *Myh6/Myh7*, northern blot, and *Mhrt779* characterization.** **a**, Quantification of cardiac *Myh6/Myh7* ratio 2–42 days after sham or TAC operation. **b**, Northern blot analysis of *Mhrt*, *Myh6* and *Myh7*. Negative: control RNA from 293T cells. Size control: 826 is recombinant *Mhrt826*; 643 (not a distinct *Mhrt* species) contains the 5' common region of *Mhrt*. Heart: adult heart ventricles. **c**, Un-cropped northern blots of *Mhrt*, *Myh6* and *Myh7*. **d**, RNA *in situ* hybridization of *Mhrt779* of adult heart ventricles. White arrowheads indicate nuclei of myocardial cells.

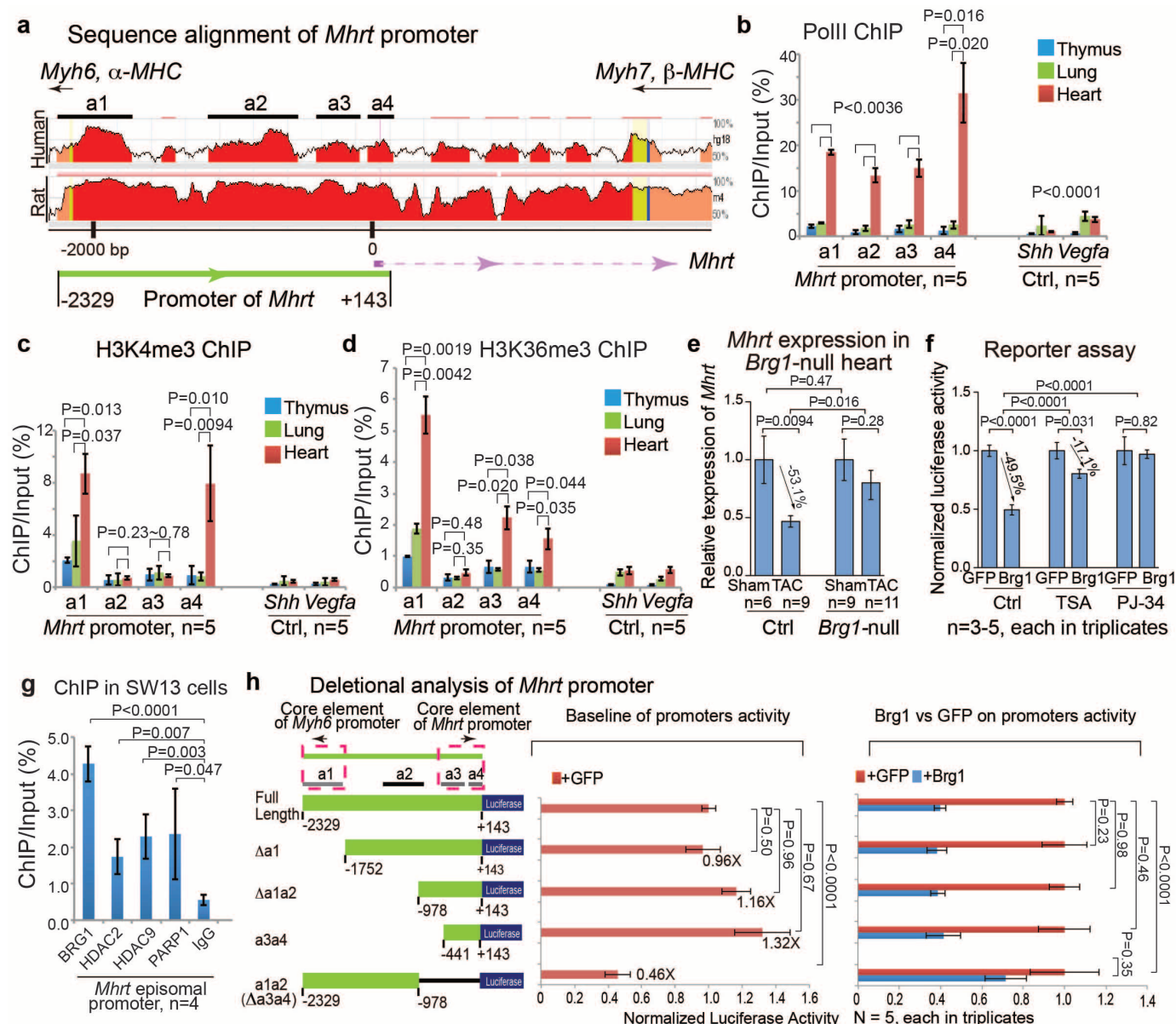
Black arrowheads indicate nuclei of endothelial, endocardial or epicardial cells. Blue: *Mhrt779*; Red: nuclear fast red. Epi, epicardium. The dashed line separates the epicardium from myocardium. Scale bars = 50  $\mu$ m. **e**, Quantification of *Tff1b*, *Hprt1*, 28S rRNA, *Neat1* and *Mhrt779* in the nuclear and cytoplasmic fraction of adult heart ventricle extracts. The nuclear/cytoplasmic ratio of *Tff1b* is set as 1. *P* values: Student's *t*-test. Error bars show s.e.m.





**Extended Data Figure 3 | Wheat germ agglutinin staining, time course and molecular marker studies of the stressed *Tg779* mice.** **a**, Wheat germ agglutinin (WGA) immunostaining 6 weeks after the sham or TAC operation. Green: WGA stain, outlining cell borders of cardiomyocytes. Blue: 4',6-diamidino-2-phenylindole (DAPI). Ctrl, control mice. Scale bars = 50  $\mu$ m. **b**, Time course of fractional shortening (FS) in control and *Tg779* mice.

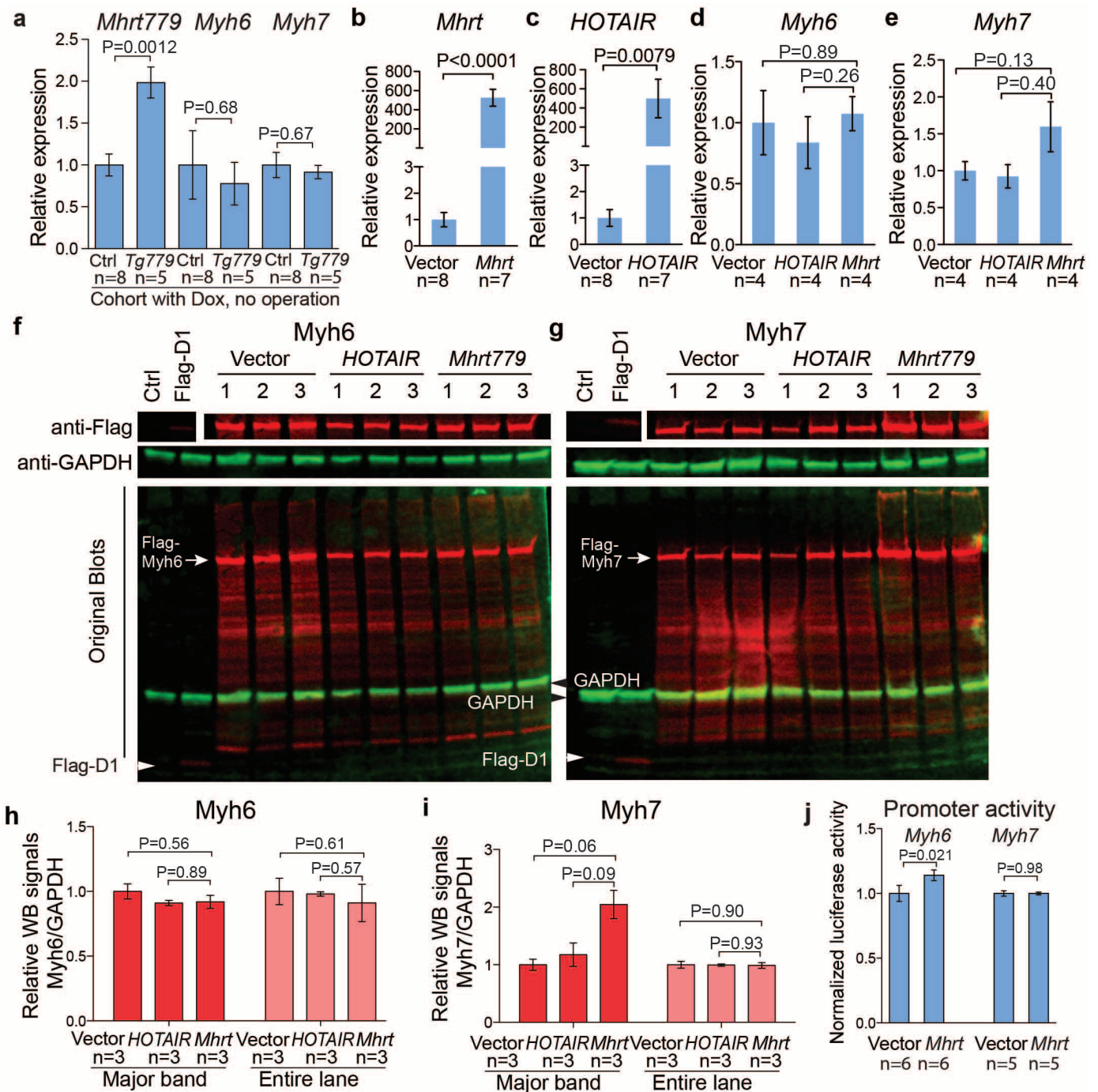
**c**, Quantification of *Anf*, *Bnp*, *Serca2* and *Tgfb1* in control and *Tg779* mice 2 weeks after sham or TAC operation. **d**, Experimental design for treatment study and time course of left ventricular fractional shortening changes. **e**, Fractional shortening of the left ventricle (LV) 8 weeks after the operation. **f**, Ventricle weight/body weight ratio of hearts harvested 8 weeks after sham or TAC operation. *P* values: Student's *t*-test. Error bars show s.e.m.



**Extended Data Figure 4 | Regulation of the *Mhrt* promoter.** **a**, Sequence alignment of *Mhrt* promoter loci from mouse, human and rat. Peak heights indicate degree of sequence homology. Black boxes (a1–a4) are sequences of high homology, which were used for further ChIP analysis. Green box region between *Myh6* and *Mhrt* is the putative *Mhrt* promoter. Red, promoter regions; salmon, introns; yellow, untranslated regions. **b–d**, ChIP–qPCR analysis of *Mhrt* promoter using antibodies against Pol II (**b**), H3K4me3 (**c**), and H3K36me3 (**d**) in tissues of adult mice. **e**, RT–qPCR quantification of *Mhrt* in

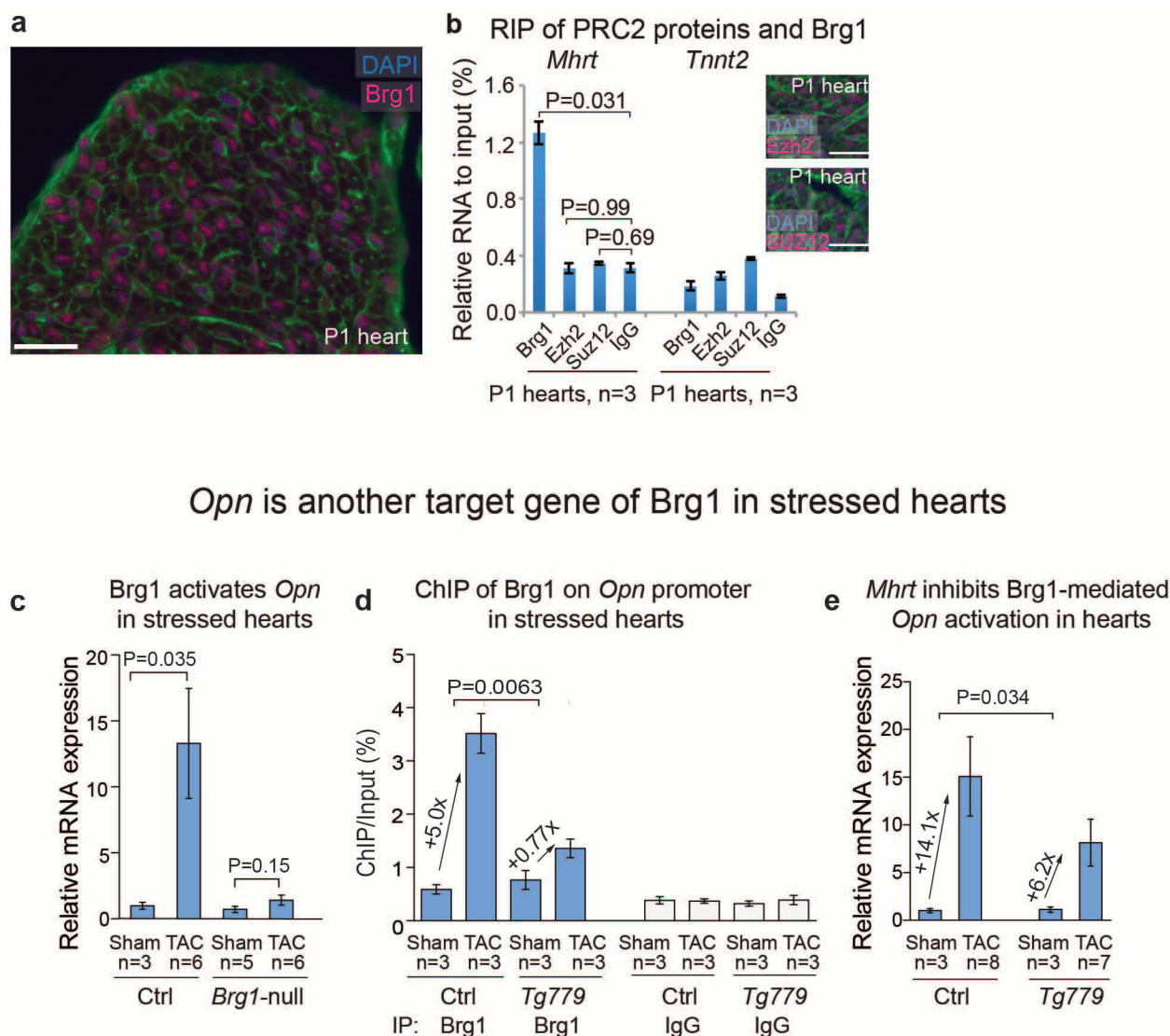
control and *Brg1*-null hearts after 7 days of TAC. Ctrl, control. *Brg1*-null, *Tnnt2-rtTA;Tre-Cre;Brg1<sup>fl/fl</sup>*. **f**, Luciferase reporter assay of *Mhrt* promoter in SW13 cells. Ctrl: dimethylsulphoxide (DMSO). PJ-34, PARP inhibitor; TSA, trichostatin (HDAC inhibitor). **g**, ChIP analysis of BRG1, HDAC2, HDAC9 and PARP1 in SW13 cells. The cells were transfected with episomal *Mhrt* promoter cloned in pREP4. **h**, Deletional analyses of the *Mhrt* promoter in luciferase reporter assays in SW13 cells. Luciferase activity of full-length *Mhrt* promoter was set up as 1. *P* values: Student's *t*-test. Error bars show s.e.m.





**Extended Data Figure 5 | *Mhrt* does not affect *Myh* expression by direct RNA sequence interference.** **a**, qPCR analysis of *Mhrt779*, *Myh6* and *Myh7* in mice without TAC operation. Expression levels were normalized to *Tfllb*, and the control is set as 1. Ctrl, control mice. **b**, **c**, RNA quantification of *Mhrt* (b) and *HOTAIR* (c) in SW13 cells transfected with Vector (pAdd2), *HOTAIR* (pAdd2-*HOTAIR*) or *Mhrt* (pAdd2-*Mhrt779*). Expression in vector-transfected cells is set as 1. Constructs containing *Myh6* or *Myh7* were co-transfected into SW13 cells used for Fig. 2b–i. **d**, **e**, RNA quantification of *Myh6* (d) and *Myh7* (e) in SW13 cells relative to *GAPDH*. **f**, **g**, Western blot

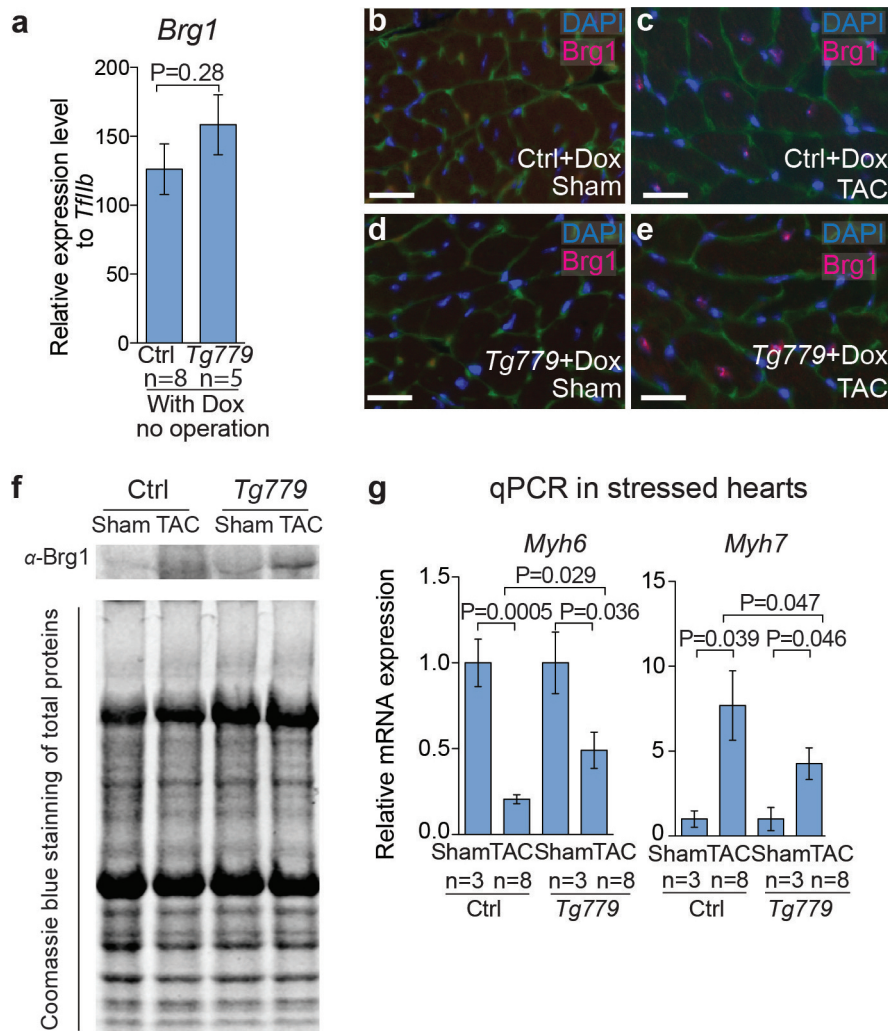
analysis of *Myh6* (f) and *Myh7* (g) in SW13 cells. Constructs containing *Myh6*- and *Myh7*-coding sequences were tagged with Flag and co-transfected with vector, *HOTAIR* or *Mhrt779*. *GAPDH* was used as the loading control. Flag-D1 was used as a positive control for the Flag antibody. **h**, **i**, Protein quantification of *Myh6* (h) and *Myh7* (i) in control and transfected SW13 cells relative to *GAPDH*. Signals of *Myh6* and *Myh7* from major bands or the entire lanes were quantified. **j**, Luciferase reporter assay of *Myh6* and *Myh7* promoters in SW13 cells transfected with vector (pAdd2) or *Mhrt* (pAdd2-*Mhrt779*). *P* values: Student's *t*-test. Error bars show s.e.m.



**Extended Data Figure 6 | RNA-IP controls; *Opn* is another target gene of Brg1 in stressed hearts.** **a**, Immunostaining of Brg1 in P1 heart. Red: Brg1. Green: WGA. Blue: DAPI. Ctrl, control. Scale bar = 50  $\mu$ m. **b**, RNA-IP of *Mhrt* in P1 hearts using antibodies against Ezh2 and Suz12. Right panels show immunostaining of Ezh2 and Suz12 in P1 hearts. PRC2, polycomb repressor complex 2. Red: Ezh2 or Suz12. Green: WGA. Blue: DAPI. Scale bars = 50  $\mu$ m.

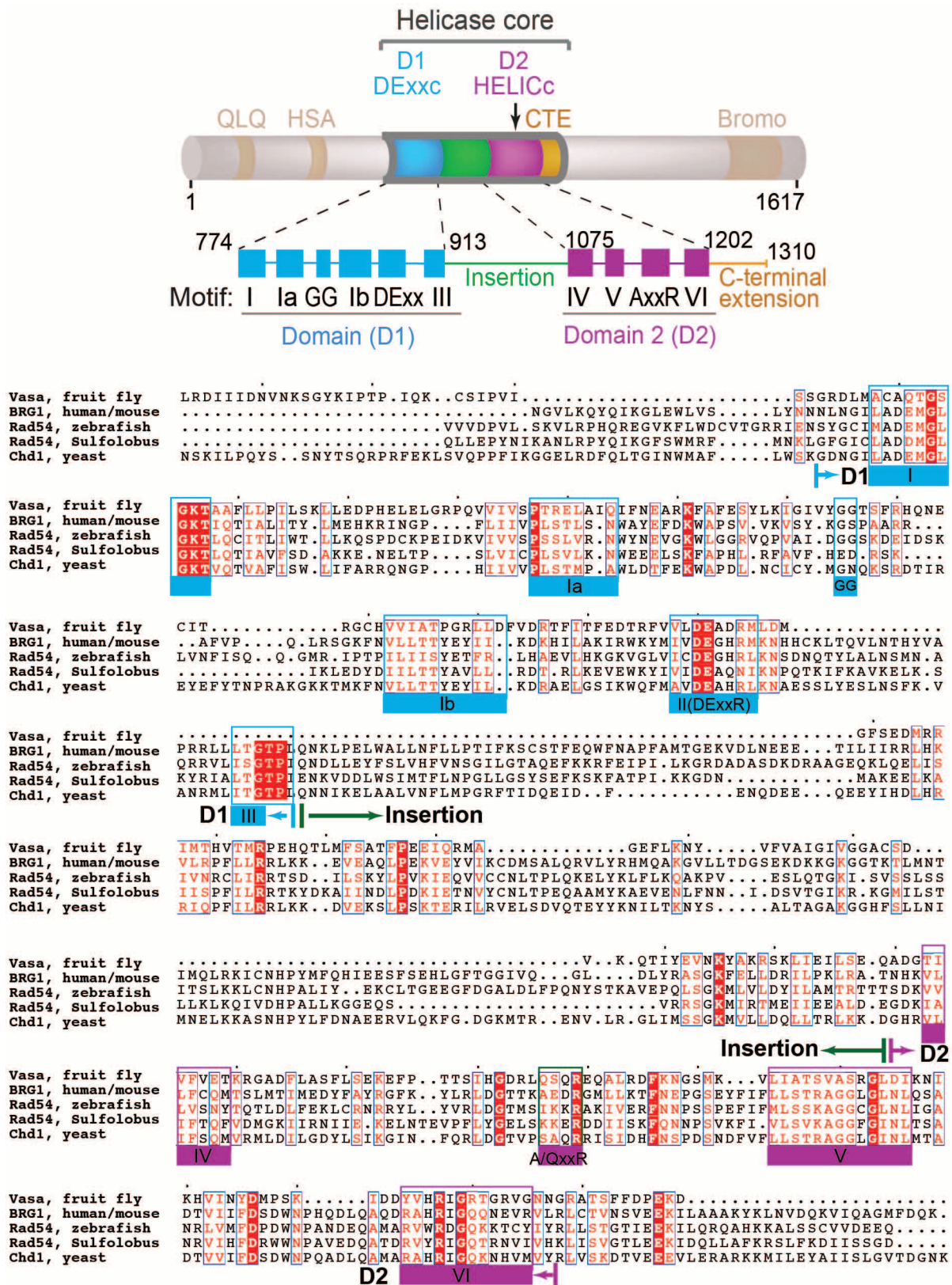
**c**, Quantification of *Opn* mRNA in control and *Brg1*-null (*Tnnt2-rtTA;Tre-Cre;Brg1<sup>fl/fl</sup>*) mice after sham or TAC operation. **d**, ChIP of Brg1 on *Opn* proximal promoter in control and transgenic (*Tg779*) mice after sham or TAC operation. **e**, Quantification of *Opn* in control and transgenic (*Tg779*) mice after sham or TAC operation. *P* values: Student's *t*-test. Error bars show s.e.m.





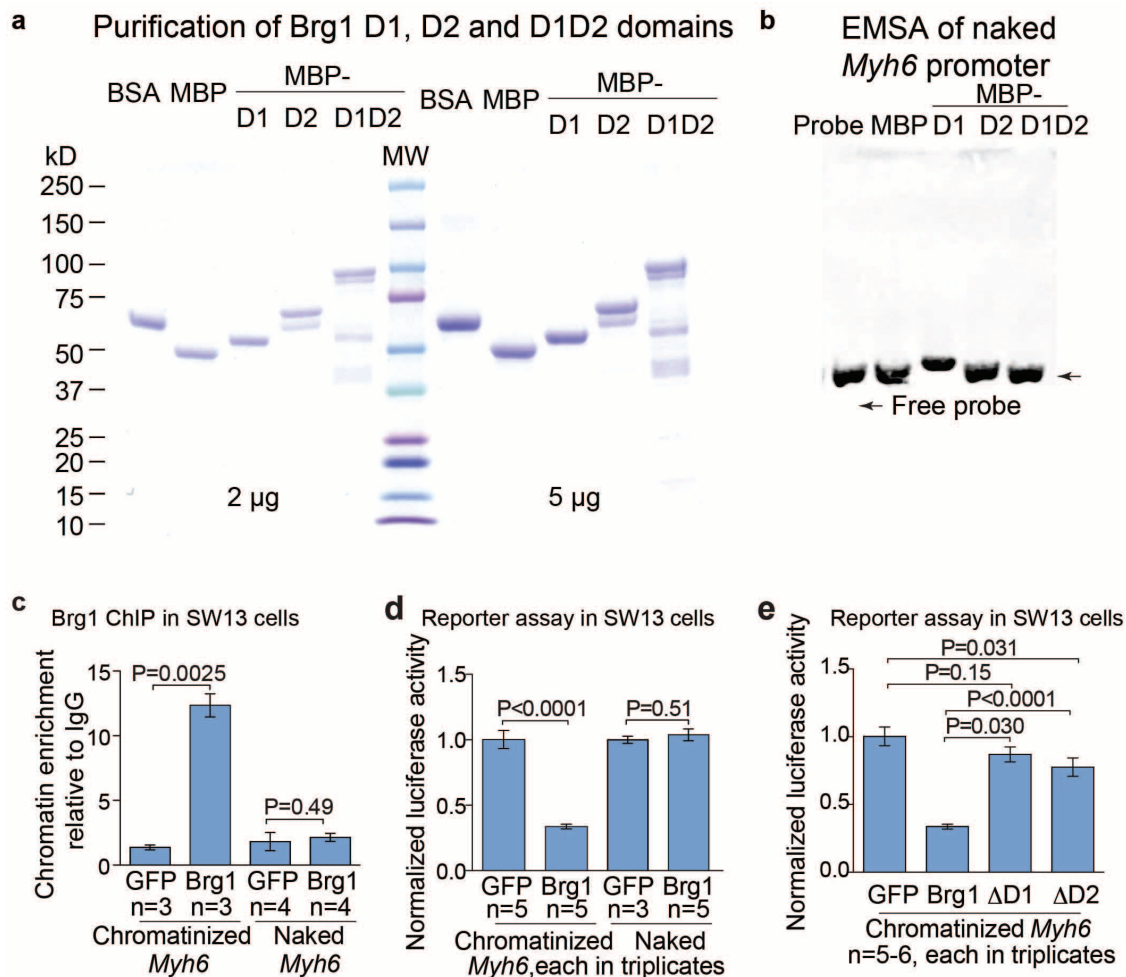
**Extended Data Figure 7 | Induction of *Mhrt779* is insufficient to change *Brg1* mRNA or protein level.** **a**, qPCR analysis of *Brg1* expression in hearts without TAC operation. Ctrl: control mice. **b–e**, Immunostaining of *Brg1* (red) in adult heart ventricles 2 weeks after sham or TAC operation. Green: WGA. Blue: DAPI. Scale bars = 50  $\mu$ m. **f**, Western blot analysis of *Brg1* and

Coomassie staining of total proteins in control or *Tg779* hearts after 2 weeks of sham or TAC operation. **g**, Quantification of *Myh6* and *Myh7* in control (Ctrl) and *Tg779* hearts after 2 weeks of sham or TAC operation. *P* values: Student's *t*-test. Error bars show s.e.m.



**Extended Data Figure 8 | Brg1 sequence alignment and motif analysis.** Schematics of the architecture of mouse Brg1 and the sequence alignment of Brg1, Vasa (fruit fly), Rad54 (zebrafish, *Sulfolobus solfataricus*) and Chd1

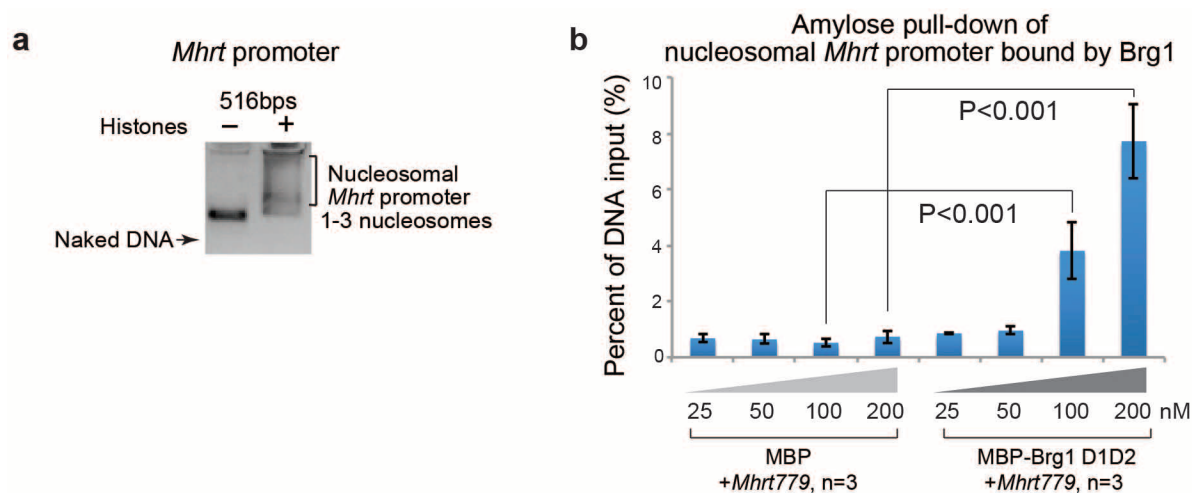
(yeast). The motifs were outlined by blue boxes (D1 domain) and purple boxes (D2 domain).



**Extended Data Figure 9 | Purification of Brg1 helicase core domains, EMSA of naked *Myh6* promoter, ChIP and reporter studies in SW13 cells.**

**a**, Coomassie blue staining of purified MBP-tagged Brg1 helicase domains. Bovine serum albumin (BSA) was loaded as a control. **b**, EMSA assay of naked *Myh6* promoter (−426 to +170) with helicase domains of Brg1. Probe: biotin-labelled *Myh6* promoter. 50 µM of MBP, MBP-D1, MBP-D2 and MBP-D1D2 proteins were used for EMSA. **c**, **d**, ChIP (c) and luciferase reporter

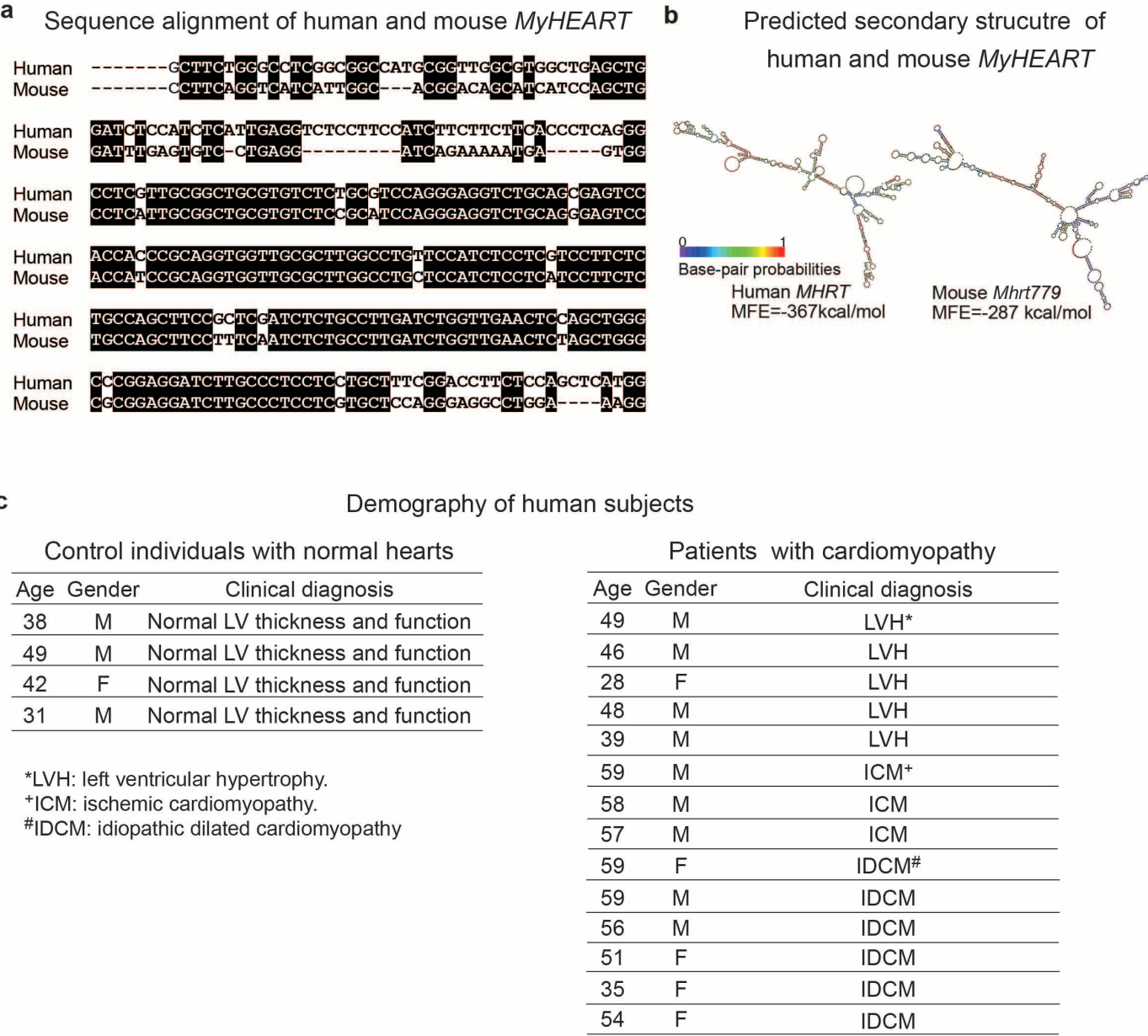
(d) analysis of Brg1 on chromatinized (episomal) and naked *Myh6* promoter in SW13 cells. GFP, green fluorescent protein control. **e**, The luciferase reporter of helicase-deficient Brg1 on chromatinized (episomal) *Myh6* promoter in SW13 cells. ΔD1: Brg1 lacking amino acids 774–913. ΔD2: Brg1 lacking amino acids 1086–1246. ChIP: H-10 antibody recognizing N terminus, non-disrupted region of Brg1. *P* values: Student's *t*-test. Error bars show s.e.m.



**Extended Data Figure 10 | Brg1 outruns *Mhrt* to bind to its target *Mhrt* promoter.** **a**, Assembly of nucleosomes on the *Mhrt* promoter (a3/4). **b**, Amylose pull-down assay: amylose was used to pull down the chromatinized

*Mhrt* promoter that was incubated with various doses of MBP and MBP-Brg1 D1D2. DNA precipitated by amylose was further quantified by qPCR. *P* values: Student's *t*-test. Error bars show s.e.m.





Extended Data Figure 11 | Sequence alignment and secondary structure prediction of human and mouse *MHRT*, and demography of heart transplantation donors. **a**, Sequence alignment of human *MHRT* and mouse *Mhrt779*. **b**, Predicted secondary structure of mouse *Mhrt779* and human

*MHRT*, using minimal free energy (MFE) calculation of RNAfold WebServer. **c**, Demography of human subjects whose tissues were used for RT-qPCR analysis (Fig. 4l). ICM, ischaemic cardiomyopathy; IDCM, idiopathic cardiomyopathy; LVH, left ventricular hypertrophy.

# DNA-damage-induced differentiation of leukaemic cells as an anti-cancer barrier

Margarida A. Santos<sup>1</sup>, Robert B. Faryabi<sup>1\*</sup>, Aysegul V. Ergen<sup>1\*</sup>, Amanda M. Day<sup>1\*</sup>, Amy Malhowski<sup>1\*</sup>, Andres Canela<sup>1</sup>, Masahiro Onozawa<sup>2</sup>, Ji-Eun Lee<sup>3</sup>, Elsa Callen<sup>1</sup>, Paula Gutierrez-Martinez<sup>4,5</sup>, Hua-Tang Chen<sup>1</sup>, Nancy Wong<sup>1</sup>, Nadia Finkel<sup>1</sup>, Aniruddha Deshpande<sup>6</sup>, Susan Sharrow<sup>7</sup>, Derrick J. Rossi<sup>4,5</sup>, Keisuke Ito<sup>8</sup>, Kai Ge<sup>3</sup>, Peter D. Aplan<sup>2</sup>, Scott A. Armstrong<sup>6</sup> & André Nussenzweig<sup>1</sup>

**Self-renewal is the hallmark feature both of normal stem cells and cancer stem cells<sup>1</sup>. Since the regenerative capacity of normal haematopoietic stem cells is limited by the accumulation of reactive oxygen species and DNA double-strand breaks<sup>2–4</sup>, we speculated that DNA damage might also constrain leukaemic self-renewal and malignant haematopoiesis. Here we show that the histone methyl-transferase MLL4, a suppressor of B-cell lymphoma<sup>5,6</sup>, is required for stem-cell activity and an aggressive form of acute myeloid leukaemia harbouring the MLL–AF9 oncogene. Deletion of *MLL4* enhances myelopoiesis and myeloid differentiation of leukaemic blasts, which protects mice from death related to acute myeloid leukaemia. MLL4 exerts its function by regulating transcriptional programs associated with the antioxidant response. Addition of reactive oxygen species scavengers or ectopic expression of FOXO3 protects *MLL4*<sup>−/−</sup> MLL–AF9 cells from DNA damage and inhibits myeloid maturation. Similar to *MLL4* deficiency, loss of *ATM* or *BRCA1* sensitizes transformed cells to differentiation, suggesting that myeloid differentiation is promoted by loss of genome integrity. Indeed, we show that restriction-enzyme-induced double-strand breaks are sufficient to induce differentiation of MLL–AF9 blasts, which requires cyclin-dependent kinase inhibitor p21<sup>Cip1</sup> (Cdkn1a) activity. In summary, we have uncovered an unexpected tumour-promoting role of genome guardians in enforcing the oncogene-induced differentiation blockade in acute myeloid leukaemia.**

Leukaemias with MLL translocations account for the majority of acute lymphoblastic leukaemias and acute myeloid leukaemias in infants, and are associated with extremely poor prognosis and response to conventional therapies<sup>7</sup>. MLL1, the founding member of the MLL family of histone methyltransferases, is essential for stem-cell self-renewal<sup>8</sup>. MLL1 fusion genes lack endogenous histone methyltransferase activity but retain MLL-associated DNA binding<sup>7,9</sup>; therefore aberrant self-renewal of myeloid progenitors and malignant cell proliferation is thought to require the recruitment of alternative histone methyltransferases to canonical MLL1 target genes<sup>7,9</sup>. In addition to MLL1, five MLL family members possess H3K4-specific methyltransferase activity. Among these, *MLL4* (also known as *Kmt2d* and orthologous to the human *MLL2* gene) has emerged as a major tumour suppressor gene but its mechanism of action and target genes are unknown<sup>5,6,10,11</sup>. To determine the role of the chromatin regulator MLL4 in normal haematopoiesis and MLL1-fusion-induced leukaemogenesis, we deleted *MLL4* in stem and progenitor cells by crossing *MLL4*<sup>fl/fl</sup> mice with transgenic mice expressing interferon-inducible *MxCre* (Extended Data Fig. 1a–d).

Total bone-marrow cellularity was equivalent in polyinosinic:polycytidylic acid (polyIC)-treated wild-type *MxCre*<sup>+</sup> and *MLL4*<sup>fl/fl</sup> *MxCre*<sup>+</sup>

mice (herein referred to as WT and *MLL4*<sup>−/−</sup> respectively) (Extended Data Fig. 1e). However, the number of Lin<sup>−</sup>, Sca1<sup>+</sup>, c-Kit<sup>+</sup> cells (LSKs), long-term haematopoietic stem cells (LT-HSCs) and myeloid (Mac1<sup>+</sup>, Gr1<sup>+</sup>) cells was significantly elevated, whereas common lymphoid progenitors and B cells were reduced in the absence of MLL4 (Extended Data Fig. 1f–j). While there was no difference in the number of myeloid-biased HSCs (Extended Data Fig. 2a, b; *P* > 0.8), there was an increased frequency of bone-marrow-derived common myeloid progenitors, and an increased myeloid colony-forming potential in the absence of MLL4 (Extended Data Fig. 2c, d). *MLL4*<sup>−/−</sup> spleens were significantly larger than controls and exhibited extramedullary haematopoiesis (Extended Data Fig. 2e–h). Thus, loss of MLL4 results in an expansion of HSCs and myeloid cells but reduced lymphopoiesis.

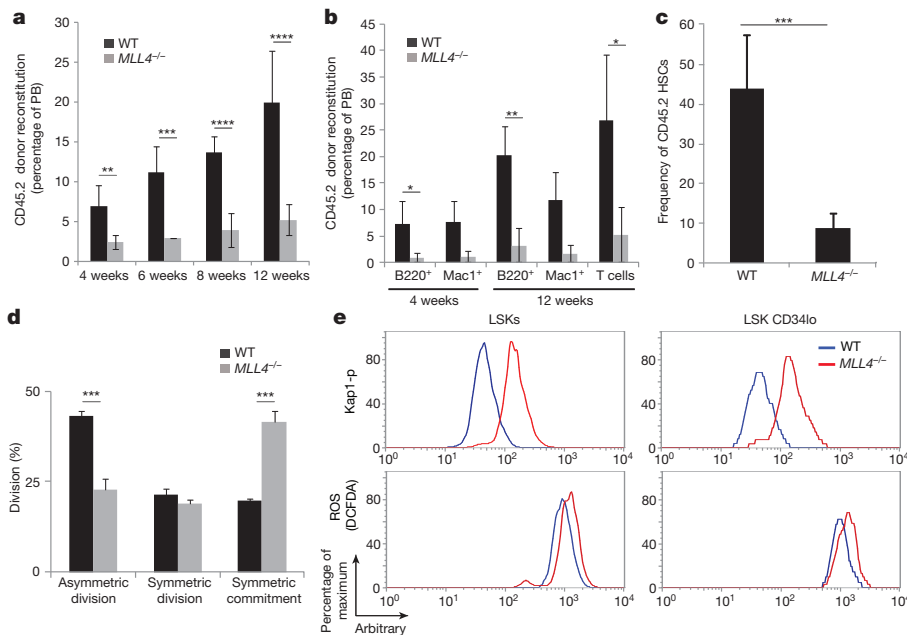
We compared the repopulating ability of a CD45.2 congenic test population (WT or *MLL4*<sup>−/−</sup> unfractionated whole bone marrow) against equal numbers of WT cells marked with the CD45.1 allele to support transplantation into lethally irradiated WT CD45.1 recipients (Extended Data Fig. 3a). Peripheral blood analysis revealed that the total CD45.2-derived reconstitution was reduced in mice transplanted with *MLL4*<sup>−/−</sup> bone marrow (Extended Data Fig. 3b–d), and *MLL4*<sup>−/−</sup> LSKs were eightfold less abundant than WT at 19 weeks after transplantation (Extended Data Fig. 3e). Competitive transplantation experiments with equal numbers of CD34<sup>lo</sup> LSKs from WT and *MLL4*<sup>−/−</sup> mice again revealed poor reconstitution of all lineages from *MLL4*<sup>−/−</sup> donor cells (Fig. 1a–c and Extended Data Fig. 3f–h), independently of any potential impact on HSC homing upon transplantation (Extended Data Fig. 3i–l). Despite these deficits in maintaining the *MLL4*<sup>−/−</sup> population in competitive assays, *MLL4*<sup>−/−</sup> bone marrow supported transplant reconstitution in non-competitive repopulation assays (Extended Data Fig. 3m–o). Thus, although MLL4 deficiency allows for haematopoietic homeostasis, the response to competitive repopulation stress is severely compromised.

Despite the fact that *MLL4*-mutant mice showed an increase in the number of HSCs and LSKs (Extended Data Fig. 1f–h), we did not detect an increase in the percentage of cycling (S/G2/M phase) cells (Extended Data Fig. 4a, b). To examine the symmetry of cell divisions during cell cycle, we purified CD150<sup>+</sup>CD48<sup>−</sup>CD41<sup>−</sup>Flt3<sup>−</sup> LSK cells and cultured them for an *in vitro* immunophenotypic division assay (Extended Data Fig. 4c)<sup>12,13</sup>. After purification, more than 90% of WT and *MLL4*<sup>−/−</sup> HSCs expressed the receptor tyrosine kinase molecule Tie2, indicative of their quiescent state (Extended Data Fig. 4d). During asymmetric division, an HSC gives rise to a copy of itself (indicated by Tie2<sup>+</sup>CD48<sup>−</sup>) and to a committed progenitor daughter cell (indicated by Tie2<sup>−</sup>CD48<sup>+</sup>) (Extended Data Fig. 4c). We found that after an initial cellular division, the frequency of asymmetric divisions was approximately twofold lower

<sup>1</sup>Laboratory of Genome Integrity, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. <sup>2</sup>The Genetics Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. <sup>3</sup>Laboratory of Endocrinology and Receptor Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. <sup>4</sup>Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, Massachusetts 02115, USA. <sup>5</sup>Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>6</sup>Human Oncology and Pathogenesis Program and Department of Pediatrics, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA.

<sup>7</sup>Experimental Immunology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. <sup>8</sup>Ruth L. and David S. Gottesman Institute for Stem Cell and Regenerative Medicine Research, Departments of Cell Biology and Medicine, Albert Einstein Cancer Center, Albert Einstein College of Medicine, Bronx, New York 10461, USA.

\*These authors contributed equally to this work.



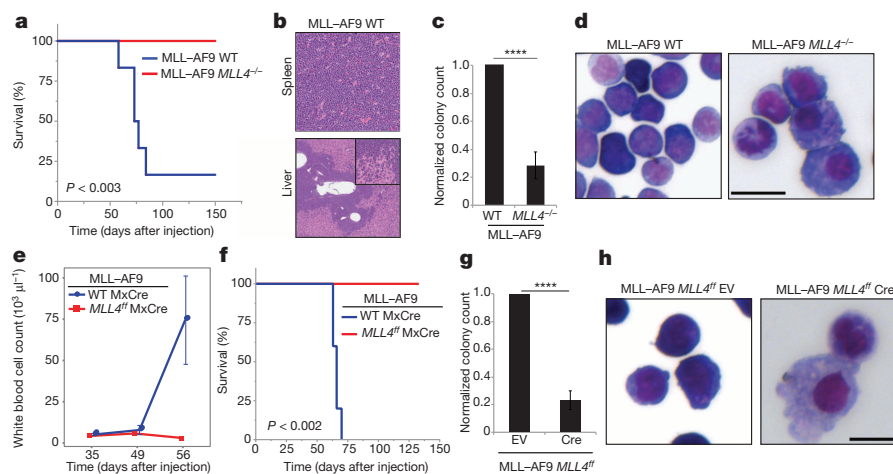
**Figure 1 | Defects in HSC function in the absence of *MLL4*.** Purified HSCs were transplanted into irradiated recipients (details in Extended Data Fig. 3f). Reconstitution levels were monitored in the peripheral blood (PB) (a, b), and the frequency of donor-derived LSK CD34<sup>lo</sup> cells was determined in the bone marrow at week 14 (c). Bar graphs show mean  $\pm$  s.d. of four and five WT and *MLL4*<sup>-/-</sup> mice, respectively. d, Division pattern of WT or *MLL4*<sup>-/-</sup> HSCs (*n* = 3 experiments, 40–77 randomly selected divisions per experiment, Extended Data Fig. 4c, d for details). e, LSK and LSK CD34<sup>lo</sup> cells from WT or *MLL4*<sup>-/-</sup> mice were stained with an antibody against phosphorylated Kap1 (top) and DCFDA (2',7'-dichlorofluorescein diacetate; bottom). (\**P* < 0.05; \*\**P* < 0.01; \*\*\**P* < 0.005; \*\*\*\**P* < 0.0001).

in *MLL4*<sup>-/-</sup> compared with WT, with a concomitant increase in the frequency of symmetric commitment (Fig. 1d). Thus, loss of *MLL4* is associated with a skewing towards symmetric commitment, which has been linked with attenuated self-renewal capacity<sup>12,13</sup>. Altogether, our data suggest that under homeostatic conditions loss of *MLL4* leads to an increase in HSCs. However, when the cells are forced to enter into cycle under conditions of stress, as during the *in vivo* repopulation or *in vitro* cell division assay, their stem-cell capacity is impaired.

To understand how *MLL4* regulates stem-cell function, we performed global analysis of gene expression changes in LSK cells. This analysis revealed that genes positively regulated by *MLL4* were associated with several processes involved in cellular response to stress (Extended Data Fig. 4e). Specifically, gene set enrichment analysis (GSEA) indicated significant enrichment of the glutathione detoxification pathway in the *MLL4* positively regulated genes (Extended Data Fig. 4f, g; false discovery rate (FDR) < 0.1), which was confirmed by quantitative real-time

reverse-transcription PCR (RT-qPCR) (Extended Data Fig. 4h). The members of the FoxO transcription factors family FoxO1, 3 and 4 (FoxOs) are also important mediators of HSC resistance to reactive oxygen species (ROS)<sup>4,14</sup>. Genes that were downregulated in FoxO-deficient LSKs were also significantly enriched among those genes downregulated in the absence of *MLL4* (FDR < 0.1, Extended Data Fig. 4i). Thus, *MLL4* deficiency in the HSC compartment deregulated the expression of genes mediating resistance to oxidative stress.

Oxidative stress and DNA damage limit HSC functional capacity<sup>2–4</sup>. Flow cytometric analysis revealed that *MLL4*<sup>-/-</sup> LSKs and HSCs (LSK CD34<sup>lo</sup>) exhibited an increase in ROS (Fig. 1e, bottom, and Extended Data Fig. 4j) and DNA damage, as measured by levels of phosphorylated Kap1 (Fig. 1e, top), a target of the DNA-damage kinases ATM, ATR and DNA-PKcs. Thus, the loss of HSC reconstitution potential and self-renewal defects in *MLL4*<sup>-/-</sup> mice are associated with the accumulation of endogenous ROS and DNA damage signalling.



**Figure 2 | *MLL4* is required for MLL-AF9-induced leukaemia.** a, Kaplan-Meier survival plots of mice injected with WT or *MLL4*<sup>-/-</sup> cells transformed with MLL-AF9 (details in Extended Data Fig. 5a). WT (*n* = 7), *MLL4*<sup>-/-</sup> (*n* = 6). b, Haematoxylin and eosin stain of WT mice at the time of death showing leukaemic cells in the spleen (top  $\times 20$ ) and liver (bottom  $\times 40$ , zoomed-in for details). Data representative of at least three mice. c, Blast colony count and colony morphology in methylcellulose, May-Grünwald-Giemsa stain (d) of WT or *MLL4*<sup>-/-</sup> cells 10 days after MLL-AF9 transformation. Bar graph shows mean  $\pm$  s.d. of four independent experiments and (d) is representative of

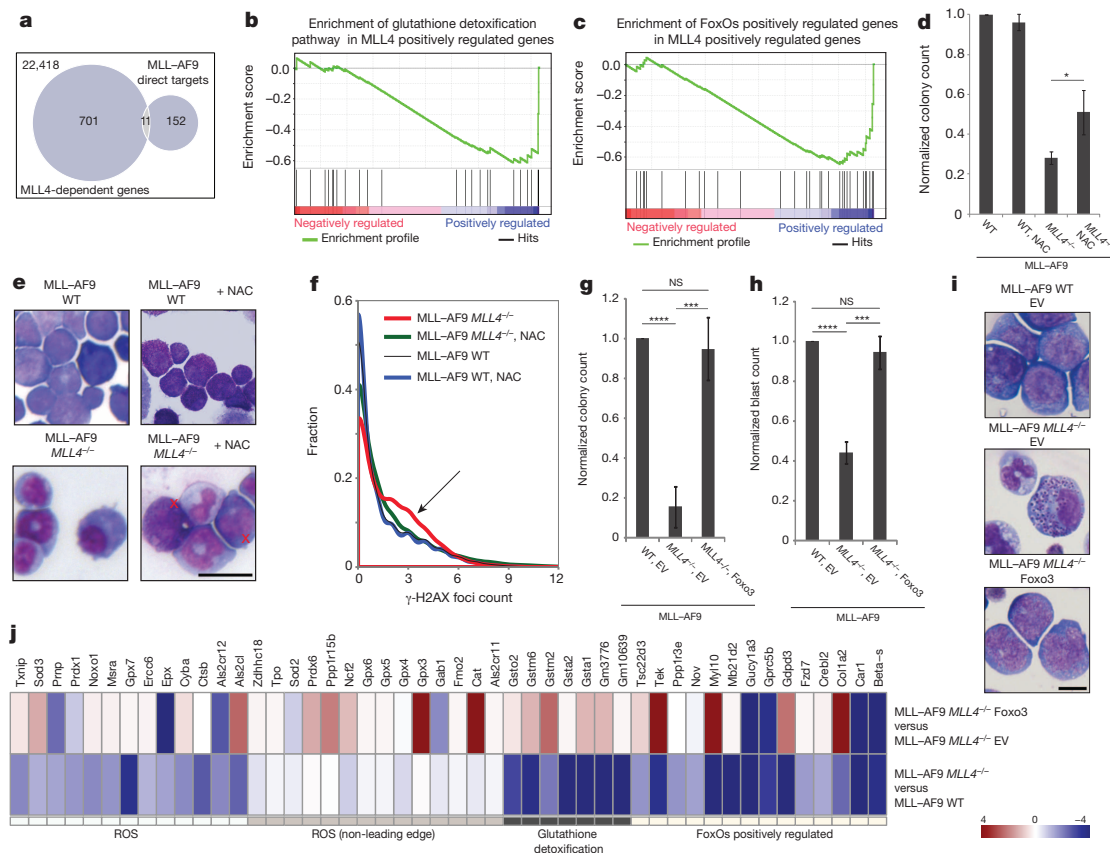
three experiments. e, Peripheral blood white cell counts (mean  $\pm$  s.d. of five mice) and (f) Kaplan-Meier survival plots of mice injected with MLL-AF9 WT MxCre or MLL-AF9 *MLL4*<sup>fl</sup> MxCre cells and subsequently treated with polyIC injections (details in Extended Data Fig. 5b). Blast colony counts (g) (mean  $\pm$  s.d. of four independent experiments) and colony morphology (h) of EV or CRE-infected *MLL4*<sup>fl</sup> MLL-AF9 cells 10 days after sort (details in Extended Data Fig. 5b; images representative of three independent experiments); EV, empty vector. Scale bars, 10  $\mu$ m.

Despite their well-established tumour suppressor functions, FoxOs are required for the maintenance of leukaemic initiating cells in a model that mimics acute myeloid leukaemias containing a translocation between *MLL1* and *AF9* genes<sup>15</sup>. To determine whether *MLL4* modifies *MLL*-*AF9* leukaemia, we introduced *MLL*-*AF9* into WT and *MLL4*<sup>-/-</sup> bone-marrow haematopoietic stem/progenitor cells with a retrovirus marked with green fluorescent protein (GFP) (*MLL*-*AF9*-IRES-GFP)<sup>16</sup> (Extended Data Fig. 5a, c). When injected into sublethally irradiated recipients, WT cells transformed with *MLL*-*AF9* caused leukaemia, with 70% of the animals succumbing by 80 days (Fig. 2a, b and Extended Data Fig. 5d). In contrast, *MLL4*-deficient cells transformed with *MLL*-*AF9* failed to cause leukaemia (Fig. 2a), even when *MLL4* was excised after cells transformed with *MLL*-*AF9* were injected into syngeneic recipients (Extended Data Fig. 5b and Fig. 2e, f); moreover, unlike non-transformed *MLL4*<sup>-/-</sup> counterparts (Extended Data Fig. 5e), *MLL*-*AF9* *MLL4*<sup>-/-</sup> cells grew *ex-vivo* more poorly than WT controls (Fig. 2c, g and Extended Data Fig. 5l), despite no detectable changes in cell death or retroviral infection frequency (Extended Data Fig. 5f, g). However, *MLL4*-deficient colonies contained fewer blasts (undifferentiated cells) and the majority of cells presented morphological characteristics associated with myeloid differentiation (Fig. 2d, h and Extended Data Fig. 5h–k, m–o). Finally, short hairpin RNA (shRNA)-mediated depletion of *MLL4* in WT *MLL*-*AF9* cells also skewed their differentiation towards myeloid

lineages in culture (Extended Data Fig. 5p, q). We conclude that although inactivating mutations of *MLL4* are found in various cancers<sup>5,6,10</sup>, *MLL4* is essential for *MLL*-*AF9*-induced leukaemia.

To identify the molecular mechanisms underlying the requirement of *MLL4* in *MLL*-*AF9* leukaemogenesis, we examined *MLL4*-dependent gene expression alterations in cells transformed with *MLL*-*AF9* (Extended Data Figs 5j and 6a). GSEA revealed a marked downregulation of genes highly expressed in the *MLL*-*AF9* leukaemic stem-cell (LSC) signature in *MLL4*-deficient cells (FDR < 0.1, Extended Data Fig. 6b), as well as global upregulation of genes that are downregulated in granulocyte macrophage progenitor-like leukaemic cells (L-GMP) and HSCs relative to committed progenitors (FDR < 0.1, Extended Data Fig. 6c, Methods). Moreover, specific markers of myeloid maturation were significantly upregulated in the absence of *MLL4* (FDR < 0.1, fold change > 2) (Extended Data Fig. 6d–f). Nevertheless, more than 93% of the *MLL*-*AF9* direct targets<sup>17</sup> (including *HOXA9* and *MEIS1*) were not differentially expressed (FDR > 0.25, Fig. 3a), suggesting an alternative mechanism by which the LSC signature and self-renewal are compromised.

Increased levels of ROS in haematopoietic progenitors are associated with myeloid differentiation<sup>4,18</sup>. On the basis of our finding that *MLL4*-deficient primary stem cells exhibit higher than normal levels of ROS, we hypothesized that myeloid differentiation of *MLL4*-deficient *MLL*-*AF9* leukaemia blasts (Fig. 2d, h) might similarly be a result of dysregulation



**Figure 3 | *MLL4* enforces the differentiation blockade in cells transformed with *MLL*-*AF9* by protecting against ROS and DNA damage.** **a–c**, RNA sequencing (RNA-seq) was performed on WT and *MLL4*<sup>fl/fl</sup> cells transformed with *MLL*-*AF9* 5 days after CRE infection as in Extended Data Fig. 5b. **a**, Fewer than 7% of *MLL*-*AF9* direct targets are deregulated in *MLL4*-deficient cells (FDR > 0.25). **b**, GSEA shows enrichment of the glutathione detoxification pathway in WT versus *MLL4*-deficient cells (FDR < 0.05) and (c) downregulation of the FOXO1/3/4 positively regulated genes in *MLL4*-deficient cells (FDR < 0.05). **d–f**, Cells transformed with *MLL*-*AF9* were expanded with or without NAC. Blast colony counts (**d**) and colony morphology (**e**) were determined (images representative of two independent experiments). Red crosses show morphological changes characteristic of blasts.

**f**, Distribution of γ-H2AX foci per cell measured by high-throughput microscopy. Arrow indicates that NAC treatment reduces the frequency of *MLL4*-deficient cells with high numbers of foci. **g**, Blast colony counts, frequency of blasts (**h**) and colony morphology (**i**) determined 10 days after sorting FOXO3- or empty-vector-expressing cells transformed with *MLL*-*AF9* (details in Extended Data Fig. 7l; images representative of three independent experiments). Bar graphs show mean ± s.d. of three independent experiments. Scale bars, 10 μm. **j**, RNA-seq profile showing that FOXO3 complementation reversed some of the *MLL4*-dependent deregulated genes included in **b** and **c**. Experiments were performed three (**g–i**) or two (**j**) independent times using the same *MLL*-*AF9* cells expressing FoxO3 or the empty vector.



of antioxidant genes. Consistent with this, the glutathione detoxification pathway and FOXO1/3/4 positively regulated genes were significantly downregulated in MLL4-deficient MLL-AF9 cells (FDR < 0.05, Fig. 3b, c). Moreover, *MLL4*<sup>-/-</sup> MLL-AF9 cells exhibited higher levels of ROS than WT cells transformed with MLL-AF9 (Extended Data Fig. 7a), accompanied by increased levels of phosphorylated Kap-1, chromosomal aberrations and  $\gamma$ -H2AX foci (Extended Data Fig. 7b–e). Thus, MLL4-deficient MLL-AF9 cells have higher than normal levels of ROS and DNA damage.

To test whether ROS was mediating the differentiation of MLL4-deficient MLL-AF9 leukaemic cells, we cultured them with antioxidants N-acetyl-L-cysteine (NAC) and catalase. As expected, *MLL4*<sup>-/-</sup> MLL-AF9 cells gave rise to significantly fewer colonies than their WT counterparts, but treatment with NAC or catalase partly reversed this phenotype (Fig. 3d and Extended Data Fig. 7f), and led to a three- to fivefold increase in the percentage of MLL4-deficient blasts (Fig. 3e and Extended Data Fig. 7g, h). This correlated with a decrease in the levels of  $\gamma$ -H2AX foci and phosphorylated Kap-1 (Fig. 3f and Extended Data Fig. 7i, j). Finally, mice that received *MLL4*<sup>f/f</sup> *MxCre*<sup>+</sup> MLL-AF9 cells and were fed daily with NAC during and after polyIC administration had reduced survival relative to animals that received the same cells and polyIC treatment but were not treated with NAC (Extended Data Fig. 7k). Together, these data suggest that protection from oxidative stress and DNA damage by MLL4 is critical to enforce the differentiation blockade and thereby promote the growth of MLL-AF9 leukaemic cells.

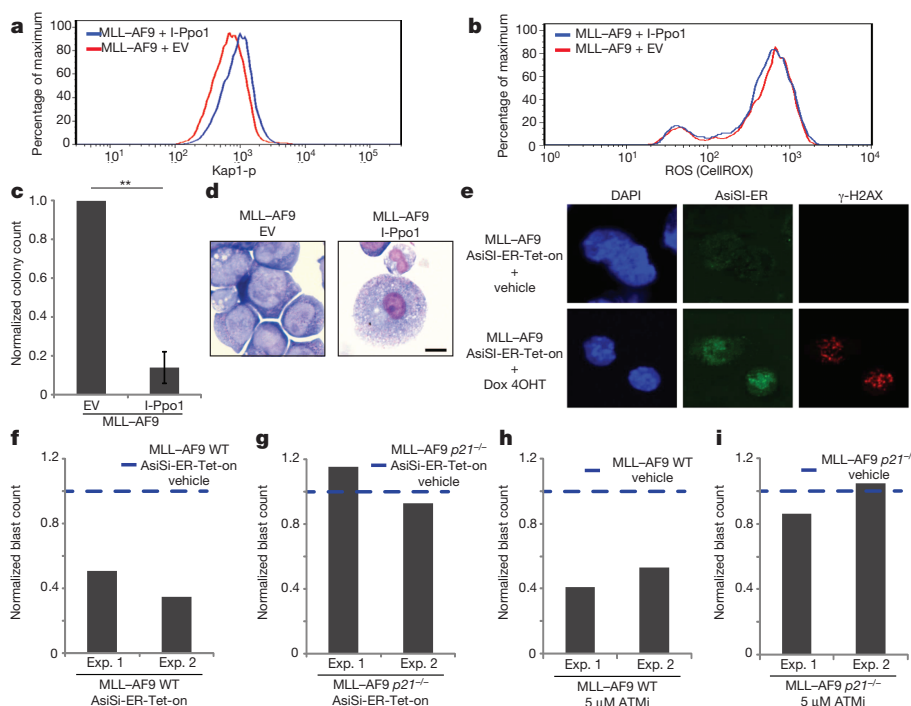
Since MLL4 positively regulates FOXO-dependent genes (Fig. 3c), we asked whether FOXO3 complementation could bypass the requirement for MLL4. Ectopic FOXO3 expression conferred nearly complete resistance to myeloid differentiation and growth impairment in *MLL4*<sup>-/-</sup> cells (Extended Data Fig. 7l and Fig. 3g–i), and a marked downregulation of MLL-AF9 LSC differentiation signature (Extended Data Fig. 7m). FOXO3 complementation also reversed many of the transcriptional defects in the glutathione, ROS and FOXO pathways (Fig. 3j), and levels of ROS were reduced in *MLL4*<sup>-/-</sup> MLL-AF9 cells overexpressing FOXO3 (Extended Data Fig. 7n). These results support the notion that the FOXO pathway is a relevant target of MLL4 in protection against ROS.

To determine whether increased oxidative stress was sufficient to confer differentiation of MLL-AF9 leukaemic blasts, we treated WT cells with hydrogen peroxide. This treatment resulted in increased ROS (Extended Data Fig. 8a), phosphorylated Kap1 (Extended Data Fig. 8b),

myeloid differentiation (Extended Data Fig. 8c) and a concomitant decrease in the frequency of blasts (Extended Data Fig. 8d, e). Thus, ROS and DNA-damage signalling are associated with myeloid differentiation.

Since genome stability is a key determinant of the ability of normal HSCs to self-renew and to sustain physiological stress<sup>23,19</sup>, we hypothesized that tumour suppressors that protect against DNA damage and oxidative stress such as ATM and BRCA1 (refs 20–22) might similarly be required to sustain the differentiation block induced by MLL-AF9. To test this, we measured the growth and morphology of WT, *BRCA1*<sup>-/-</sup> (*BRCA1*<sup>f/f</sup> × *Mx-Cre* mice treated with polyIC (Extended Data Fig. 8f)) and *ATM*<sup>-/-</sup> bone-marrow cells. In granulocyte-macrophage colony-forming unit assays (CFU-GM), loss of ATM and BRCA1 led to a small deficit in cell growth (Extended Data Fig. 8g, h). When transformed with MLL-AF9, these DNA repair mutants were incapable of maintaining *in vitro* self-renewal and proliferation without differentiation (Extended Data Fig. 8i–n). Moreover, treatment of WT MLL-AF9 cells with a specific ATM inhibitor (Ku55933, ATMi) for 48 h led to a 275% increase in their differentiation with negligible cell death (Extended Data Fig. 8o, p). Similar results were obtained with a specific ATR inhibitor (Extended Data Fig. 8q). While deficiencies in either BRCA1 or ATM led to myeloid maturation, loss of ATM was associated with increased levels of DNA damage and ROS (Extended Data Fig. 9a, b), whereas BRCA1 suppression led to DNA-damage accumulation without a detectable increase in ROS (Extended Data Fig. 9c, d). We conclude that BRCA1, ATM and ATR are critical for cytoprotective responses that maintain the differentiation blockade induced by the MLL-AF9 oncogene.

To separate the effects of ROS and DNA damage, we generated double-strand breaks (DSBs) directly in WT MLL-AF9 cells with a homing endonuclease I-Ppo1<sup>23</sup> (Extended Data Fig. 10a). I-Ppo1-infected cells exhibited an increase in levels of Kap-1 phosphorylation but no change in ROS relative to cells infected with empty vector (Fig. 4a, b). Moreover, upon re-plating we observed fivefold fewer colonies in I-Ppo1-infected cells (Fig. 4c), a fivefold reduction in the frequency of blasts (Extended Data Fig. 10b) and an induction of mature myeloid lineages (Fig. 4d). To rule out a possible toxic effect of I-Ppo1 interfering with ribosomal biogenesis due to the presence of a recognition site in the 28S ribosomal RNA genes, we used a second inducible restriction enzyme AsiSI (AsiSI-ER-Tet-on)<sup>24</sup> to generate DSBs directly in leukaemic cells (Extended Data Fig. 10c). Upon co-administration of doxycycline



**Figure 4 | DNA damage directly induces differentiation of leukaemic cells.** The levels of phosphorylated Kap1 (a) and ROS (b), blast colony counts (c) and colony morphology (d) were determined in cells transformed with MLL-AF9 expressing I-Ppo1 or an empty vector (details in Extended data Fig. 10a). Bar graph shows mean  $\pm$  s.d. of three independent experiments. d, Images were acquired once. e–g, WT or *p21*<sup>-/-</sup> cells transformed with MLL-AF9 were infected with a retrovirus containing the inducible restriction enzyme AsiSI-ER-Tet-on (details in Extended data Fig. 10c). e,  $\gamma$ -H2AX foci and AsiSI-ER (antibody to ER) staining in WT cells 24 h after treatment with doxycycline and 4OHT. Sixty-five per cent of cells were positive for  $\gamma$ -H2AX foci in WT and *p21*<sup>-/-</sup> cells (Extended data Fig. 10h; one representative of two experiments). Frequency of blasts in WT (f) and *p21*<sup>-/-</sup> (g) cells with or without the restriction enzyme. h, i, WT and *p21*<sup>-/-</sup> MLL-AF9 cells were treated with 5  $\mu$ M of ATMi or vehicle for 48 h and the frequency of blasts was determined. f–i, The frequency of blasts is normalized to vehicle-treated cells in independent experiments.

and 4-hydroxytamoxifen (4OHT), AsiSI-ER became expressed and translocated from cytoplasm to nucleus; and  $\gamma$ -H2AX was induced in 65% of the cells with no change in the levels of ROS (Fig. 4e; Extended Data Fig. 10d). By 48 h after treatment, there was 60% reduction in the frequency of blasts (Fig. 4f). Together these data suggest that DSBs either generated indirectly through the production of ROS or produced directly can bypass the MLL–AF9 oncogene-induced differentiation blockade.

Recent studies have uncovered connections between DNA-damage checkpoint pathways, stem-cell self-renewal and differentiation<sup>19,25</sup>. Since DNA-damage-induced differentiation in MLL–AF9 transformed cells was associated with impaired growth, we wondered whether the cyclin-dependent kinase inhibitor p21, which regulates the G1–S checkpoint in response to DNA damage, could be involved in terminal differentiation. In contrast to *MLL4*-, *ATM*- and *BRCA1*-deficiency, loss of *p21* had no negative impact on proliferation or myeloid differentiation in the setting of MLL–AF9 transformation (Extended Data Fig. 10 e, f). However, after exogenous DNA damage, p21 transcripts were induced (Extended Data Fig. 10g). Moreover, when exogenous DSBs were generated either by activation of AsiSI-ER (Fig. 4e–g and Extended Data Fig. 10c, h) or by ATMi treatment (Fig. 4h, i and Extended Data Fig. 10i), p21-deficient MLL–AF9 cells were resistant to DNA-damage-induced differentiation and growth inhibition. These data suggest that DNA-damage-induced cell-cycle exit and differentiation of cells transformed with MLL–AF9 are coupled by the activation of p21.

The ‘oncogene-induced replication stress’ model for cancer development posits that DNA damage induced by oncogenes in pre-cancerous lesions activates ATM and p53, which in turn trigger cell-cycle arrest, senescence and apoptosis<sup>26</sup>. These well-established DNA-damage checkpoints raise the barrier against tumour progression, but at an advanced disease state this barrier is breached by mutations in *ATM* and *p53*, which promote genome instability and cancer<sup>26</sup>. In contrast, our results argue that DNA-damage response proteins are activated in response to MLL-fusion oncogenes, but in this case they are required for tumorigenic function (Extended Data Fig. 10j). In line with this, suppression of the ATR kinase inhibits acute myeloid leukaemia driven by the MLL–ENL oncogene<sup>27</sup>. One potential mechanism by which DNA damage can induce myeloid differentiation is by lengthening the cell cycle. Indeed, recent studies showed that retroviral transduction of p21 in lymphomyeloid progenitors induced cell-cycle lengthening and consequent accumulation of the lineage determining PU.1 transcription factor, which favours macrophage differentiation<sup>28</sup>. Similarly, we hypothesize that when MLL–AF9 oncogene-induced DNA damage in leukaemic cells reaches beyond a certain threshold, p21 is activated, cells exit the cycle and initiate terminal differentiation. As a corollary, DNA repair pathway inhibitors, such as ATMi/ATRi described here (Extended Data Fig. 8o, q), may prove to be a promising modality of differentiation therapy for the treatment of MLL-associated leukaemia.

## METHODS SUMMARY

*MLL4<sup>fl/fl</sup>* mice were crossed to *Mx1-Cre* (The Jackson Laboratory) mice. Deletion of *MLL4* was achieved by injecting *MLL4 Mx1-Cre* mice with 300  $\mu$ g of polyIC five times every other day. The experiments were performed 3 weeks after the last polyIC injection. Retroviruses were used to infect bone-marrow cells harvested 4 days after administration of fluorouracil (5-FU) (250 mg per kg). After infection, cells were maintained in methylcellulose in the presence of stem cell factor (SCF; 100 ng ml<sup>-1</sup>), IL3 (10 ng ml<sup>-1</sup>) and IL6 (10 ng ml<sup>-1</sup>) and used for *in vitro* assays or injected into irradiated recipients.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 13 August 2013; accepted 13 May 2014.

Published online 27 July 2014.

1. Magee, J. A., Piskounova, E. & Morrison, S. J. Cancer stem cells: impact, heterogeneity, and uncertainty. *Cancer Cell* **21**, 283–296 (2012).

2. Nijink, A. *et al.* DNA repair is limiting for haematopoietic stem cells during ageing. *Nature* **447**, 686–690 (2007).
3. Rossi, D. J. *et al.* Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. *Nature* **447**, 725–729 (2007).
4. Tothova, Z. *et al.* FoxOs are critical mediators of hematopoietic stem cell resistance to physiologic oxidative stress. *Cell* **128**, 325–339 (2007).
5. Morin, R. D. *et al.* Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298–303 (2011).
6. Pasqualucci, L. *et al.* Analysis of the coding genome of diffuse large B-cell lymphoma. *Nature Genet.* **43**, 830–837 (2011).
7. Krivtsov, A. V. & Armstrong, S. A. MLL translocations, histone modifications and leukaemia stem-cell development. *Nature Rev. Cancer* **7**, 823–833 (2007).
8. Jude, C. D. *et al.* Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. *Cell Stem Cell* **1**, 324–337 (2007).
9. Somervaille, T. C. & Cleary, M. L. Grist for the MLL: how do MLL oncogenic fusion proteins generate leukemia stem cells? *Int. J. Hematol.* **91**, 735–741 (2010).
10. Lawrence, M. S. *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**, 495–501 (2014).
11. Lee, J. E. *et al.* H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *eLife* **2**, e01503 (2013).
12. Ito, K. *et al.* A PML–PPAR- $\delta$  pathway for fatty acid oxidation regulates hematopoietic stem cell maintenance. *Nature Med.* **18**, 1350–1358 (2012).
13. Yamamoto, R. *et al.* Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell* **154**, 1112–1126 (2013).
14. Kops, G. J. *et al.* Forkhead transcription factor FOXO3a protects quiescent cells from oxidative stress. *Nature* **419**, 316–321 (2002).
15. Sykes, S. M. *et al.* AKT/FOXO signaling enforces reversible differentiation blockade in myeloid leukemias. *Cell* **146**, 697–708 (2011).
16. Krivtsov, A. V. *et al.* Transformation from committed progenitor to leukaemia stem cell initiated by MLL–AF9. *Nature* **442**, 818–822 (2006).
17. Bernt, K. M. *et al.* MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. *Cancer Cell* **20**, 66–78 (2011).
18. Owusu-Ansah, E. & Banerjee, U. Reactive oxygen species prime *Drosophila* haematopoietic progenitors for differentiation. *Nature* **461**, 537–541 (2009).
19. Wang, J. *et al.* A differentiation checkpoint limits hematopoietic stem cell self-renewal in response to DNA damage. *Cell* **148**, 1001–1014 (2012).
20. Bae, I. *et al.* BRCA1 induces antioxidant gene expression and resistance to oxidative stress. *Cancer Res.* **64**, 7893–7909 (2004).
21. Gorrini, C. *et al.* BRCA1 interacts with Nrf2 to regulate antioxidant signaling and cell survival. *J. Exp. Med.* **210**, 1529–1544 (2013).
22. Ito, K. *et al.* Reactive oxygen species act through p38 MAPK to limit the lifespan of hematopoietic stem cells. *Nature Med.* **12**, 446–451 (2006).
23. Argast, G. M., Stephens, K. M., Emond, M. J. & Monnat, R. J. Jr. I-Pol and I-Crel homing site sequence degeneracy determined by random mutagenesis and sequential *in vitro* enrichment. *J. Mol. Biol.* **280**, 345–353 (1998).
24. Iacovoni, J. S. *et al.* High-resolution profiling of  $\gamma$ H2AX around DNA double strand breaks in the mammalian genome. *EMBO J.* **29**, 1446–1457 (2010).
25. Inomata, K. *et al.* Genotoxic stress abrogates renewal of melanocyte stem cells by triggering their differentiation. *Cell* **137**, 1088–1099 (2009).
26. Halazonetis, T. D., Gorgoulis, V. G. & Bartek, J. An oncogene-induced DNA damage model for cancer development. *Science* **319**, 1352–1355 (2008).
27. Schoppy, D. W. *et al.* Oncogenic stress sensitizes murine cancers to hypomorphic suppression of ATR. *J. Clin. Invest.* **122**, 241–252 (2012).
28. Kueh, H. Y., Champhekar, A., Nutt, S. L., Elowitz, M. B. & Rothenberg, E. V. Positive feedback between PU.1 and the cell cycle controls myeloid differentiation. *Science* **341**, 670–673 (2013).

**Acknowledgements** We thank all members of the A. Nussenzweig laboratory, J. Daniel and P. Love for discussions; L. Granger for flow cytometry; M. J. Kruhlak for microscopy; K. Zhao for RNA sequencing; R. Anderson and K. Smith for animal care; K. Naka for the FOXO3 retrovirus; G. Legube for the pBABE-AsiSI-ER plasmid; J. Zuber for pLEPG and pRT3GEPIR plasmids; O. Fernandez-Capetillo for ATRi; and S. John for suggestions. S.A.A. was supported by the Leukemia and Lymphoma Society and National Cancer Institute grants CA66996 and CA140575. This work was supported by the Intramural Research Program of the National Institutes of Health, the National Cancer Institute and the Center for Cancer Research, and an Ellison Medical Foundation Senior Scholar in Aging Award to A.N.

**Author Contributions** M.A.S., R.B.F., P.D.A., S.A.A. and A.N. participated in the study design. M.A.S., A.V.E., A.M.D., A.M., N.F., H.C. and N.W. performed mouse breeding, HSC analysis, transplantation and leukaemia experiments, and analysed data. R.B.F. performed computational experiments. E.C. led the genome stability experiments and analysed data. A.C. generated and performed experiments with AsiSI-ER-Tet-on and MLL4 shRNA and qPCR. P.G.-M. and D.J.R. supervised HSC experiments and performed serial colony assays; S.S. supervised flow cytometry. J.-E.L. and K.G. generated targeting vector and MLL4-deficient mice. K.I. performed the *in vitro* immunophenotypic division assay. M.O. quantified colony morphology in cytopins. A.D. generated and tested MLL–AF9 vectors. M.A.S. and A.N. wrote the manuscript and all authors reviewed it. A.N. supervised the project.

**Author Information** Data sets have been deposited in the Gene Expression Omnibus under accession number GSE57147. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.N. ([andre\\_nussenzweig@nih.gov](mailto:andre_nussenzweig@nih.gov)).

## METHODS

**Mice.** *MLL4<sup>f/f</sup>* mice were generated by inserting a loxp-FRT-Neo-FRT cassette 5' to exon 19 and a loxp site 3' of exon 16 to generate the targeting vector. Targeted embryonic stem cells gave germline transmission and the neo cassette was removed by crossing the mice with Flp deleter mice. Cre recombination removed exons 16–19 and generated a frameshift mutation that resulted in a truncated protein missing the carboxy (C)-terminal ~4,200 amino acids<sup>11</sup>. These animals were maintained in a B6/129 background. *BRCA1<sup>f(A11)/f(A11)</sup>* (NCI mouse repository), B6-Ly5.2/Cr (NCI mouse repository), B6;129S2-*Cdkn1a<sup>tm1Tyj/J</sup>* (*p21<sup>-/-</sup>*) (The Jackson Laboratory), *Mx1-Cre* (The Jackson Laboratory) and *ATM<sup>-/-</sup>* (provided by A. Wynshaw-Boris) mice have been described. Experiments were performed with 6- to 10-week-old mice. Males and females were equally distributed between different experimental groups. All mice were housed in the Frederick National Laboratory and treated with procedures approved by the National Institutes of Health Animal Care and Use Committee.

**PolyIC treatment.** Three hundred micrograms of polyIC (Sigma Aldrich) was administered by intraperitoneal injection five times every other day. The experiments were performed 3 weeks after the last polyIC injection.

**Isolation of bone-marrow cells, flow cytometry and HSC isolation.** Bone-marrow cells were flushed from the long bones (tibia and femurs) and stained in PBS (Corning Cellgro) supplemented with 2% of inactivated fetal bovine serum (Gemini BioProducts) with the following antibodies from BD Biosciences: B220 FITC, CD11b FITC, PE, APC or APCy7, CD11c FITC, PE, APC or APC-cy7, CD4 FITC, PE or APC, CD8 FITC, PE or APC, NK1.1 FITC, Ter119 FITC, CD3 FITC (when gates on lineage negative bone-marrow cells were used, FITC-conjugated antibodies were used), c-Kit PE or APC, Flk2 PE, CD45.2 biotin or FITC, CD45.1 APC, IL7R Alexa 647, CD34 APC, CD150 PE. From eBiosciences: Sca-1 PEcy7, AA4.1 bio or PEcy-7, streptavidin APC. From Invitrogen: streptavidin Pacific blue. DAPI was used to exclude dead cells. Flow cytometry was performed on a fluorescence-activated cell sorting (FACS) Calibur, LSRII or LSR Fortessa (BD Biosciences).

**Long-term repopulation assays.** Two hundred sort-purified LSK bone-marrow cells harvested from 6- to 10-week-old mice were mixed with 500,000 congenic whole bone-marrow cells and injected intravenously into lethally irradiated (900 rad) recipients. For whole bone-marrow transplants, bone-marrow cells ( $4 \times 10^6$  to  $8 \times 10^6$ ) harvested from 6- to 10-week-old mice were injected intravenously into lethally irradiated (900 rad) recipients. Beginning 4 weeks after transplantation and continuing for 12–19 weeks, blood was obtained from the tail veins of recipient mice, subjected to lysis of red blood cells (ACK lysing buffer, Quality Biological) and stained with antibodies to monitor engraftment.

**HSC cell division assays.** Division patterns of HSCs were determined as previously described<sup>12</sup>. Briefly, single-cell deposition of bone-marrow CD150<sup>+</sup>CD48<sup>-</sup>CD41<sup>-</sup>Flt3<sup>-</sup>CD34<sup>-</sup>KSL cells from WT or *MLL4<sup>-/-</sup>* mice (total 90–180 cells per experiment) were cultured in StemSPAN (StemCell Technologies) supplemented with 50 ng ml<sup>-1</sup> SCF and 50 ng ml<sup>-1</sup> thrombopoietin. Cells were stained with anti-mouse Tie2 antibody, anti-mouse CD48 antibody and DAPI at day 3. Student's *t*-test was used to determine statistical significance, and statistical significances in division pattern and the disturbance of division pattern were also confirmed by log-linear model and  $\chi^2$  test (data not shown).

**HSC and progenitor replating and granulocyte-macrophage colony-forming unit assays.** For HSC isolation from whole bone marrow, c-Kit enrichment was performed using CD117 magnetic beads (Miltenyi). Cells were stained with antibodies against lineage (CD3, CD4, CD8, B220, Ter119, Mac1, Gr1 and Il7ra), Sca1, c-Kit, CD34, Flk2, CD150 and Pl. HSCs were sorted as Pl<sup>-</sup>, Lin<sup>-</sup>, c-Kit<sup>+</sup>, Sca1<sup>+</sup>, CD34<sup>-</sup>, Flk2<sup>-</sup>, CD150<sup>+</sup> on a FACS Aria II (Becton, Dickinson) before genomic DNA isolation and PCR quantification of deletion efficiency. For *in vitro* replating assays, 1,000 myeloid progenitors from WT and *MLL4<sup>-/-</sup>* mice were FACS sorted from bone marrow, plated in triplicate in MethoCult M3434 (StemCell Technologies), and grown at low (5%) oxygen conditions. After primary quantification, triplicates were pooled and counted. 10,000 cells per well were plated in triplicates for subsequent analysis. All colonies were quantified after 10 days of growth. For granulocyte-macrophage colony-forming unit assays, whole bone-marrow cells were plated in MethoCult M3434 (StemCell Technologies) in the presence of SCF (100 ng ml<sup>-1</sup>), IL3 (10 ng ml<sup>-1</sup>) and IL6 (10 ng ml<sup>-1</sup>) (all from Peprotech). Colonies were scored between days 10 and 12.

**Intracellular staining of phosphorylated Kap1 and detection of ROS.** To detect Kap1-p cells were fixed and permeabilized using the BD Cytofix/Cytoperm Kit (BD Biosciences) as described by the manufacturer. Anti Kap1-p (Bethyl) was added to the cells for 1 h at 4 °C followed by the secondary antibody (either mouse-anti-rabbit Alexa Fluor 488 or Pacific Blue (Invitrogen)). To detect ROS, cells were incubated with DCFDA (2',7'-dichlorofluorescein diacetate; Invitrogen) or CellROX Deep Red Reagent (Invitrogen) according to the manufacturer's instructions for 30 min at 37 °C followed by flow cytometry.

**Immunofluorescence, metaphase analysis and microscopy.** For immunofluorescent staining of  $\gamma$ -H2AX and Asisi-ER, the ER $\alpha$  Antibody (sc-543; Santa Cruz; 1/500) and anti-phospho-Histone H2A.X (Ser139) (JBW301 Millipore, 1/5,000) were used. Cells were treated with doxocycline at 1  $\mu$ g ml<sup>-1</sup> and 4OHT at 1  $\mu$ M for 24 and 4 h respectively, before fixation and processing as described<sup>29</sup>. Cells were harvested for metaphase analysis as described<sup>30</sup>. Imaging of  $\gamma$ -H2AX foci was performed using a wide-field epi-fluorescence Zeiss Axio Observer Z1 microscope equipped with a  $\times 20$  plan apochromatic lens (numerical aperture 0.8), motorized stage and Zeiss AxioCam CCD (charge-coupled device) camera. Images were acquired and processed using Zeiss Zen imaging software with a custom-made algorithm for foci detection and then filtered based on the nucleus area, staining background and cell morphology. For acquisition of May–Grunwald–Giemsa stained cytospin slides, images were collected using Zeiss Zen image acquisition software controlling an AxioObserver Z1 wide-field microscope equipped with a plan-apochromat  $\times 63$  (numerical aperture 1.4) objective lens and an AxioCam MRC5 colour CCD camera.

**Plasmids, transformation and culture of murine cells and generation of leukaemias *in vivo*.** The following plasmids were provided by S. Armstrong: MSCV-MLL-AF9-IRES-GFP, MSCV-MLL-AF9-neo, MSCV-Cre-IRES-Tomato Red and MSCV-IRES-Tomato Red. For construction of the Asisi-ER-Tet-on vector, a fragment containing HA-ER-Asisi was PCR amplified from pBabe-Asisi-ER and cloned under the control of TRE3G doxycycline-inducible promoter of pRT3GEPiR; BamHI-MluI sites were used to remove GFP-miR-E, then HA-ER-Asisi was cloned in pRT3GEPiR. pBabe-Asisi-ER and pRT3GEPiR were gifts from G. Legube and J. Zuber, respectively. Asisi-ER-Tet-on and FoxO3 (FoxO3-IRES-GFP) retroviruses were generated in 293T cells. Retroviruses were used to infect bone-marrow cells harvested 4 days after administration of 5-FU (250 mg per kg) as previously described<sup>31</sup>. After infection, cells were maintained in methylcellulose (Methocult, StemCell Technologies) in the presence of SCF (100 ng ml<sup>-1</sup>), IL3 (10 ng ml<sup>-1</sup>) and IL6 (10 ng ml<sup>-1</sup>) (all from Peprotech). Eight days after infection,  $3 \times 10^6$  cells were injected intravenously into sublethally irradiated recipients (650 rad). White blood-cell counts were monitored from peripheral blood collections using a Hemavet (Drew Scientific). For all colony assays, MLL-AF9-infected cells were plated at 1,000 or 5,000 cells in Methocult in the presence of IL3, IL6 and SCF as above. When used, NAC (Sigma Aldrich) was added at 1  $\mu$ M, Catalase (Sigma Aldrich) at 100  $\mu$ g ml<sup>-1</sup> and ATMi KU55933 (TOCRIS Bioscience) at 5  $\mu$ M. ATRi<sup>32</sup> was used at 1  $\mu$ M. For *in vivo* experiments, 5 mg ml<sup>-1</sup> of NAC was added to the drinking water.

**Colony assays and cell morphology staining.** For colony assays, cells were plated in M3434 cytokine-enriched methylcellulose (Stem Cell Technologies) according to the manufacturer's instructions. Cytospins were performed in a Shandon Cytospin 4 (Thermo Scientific) and cells were stained first with May–Grunwald dye and then Giemsa stain (both from Sigma Aldrich). For the experiments with MLL-AF9 cells stably infected with Asisi-ER-Tet-on, Asisi-ER was induced with doxocycline at 1  $\mu$ g ml<sup>-1</sup> and 4OHT at 1  $\mu$ M for 24 and 4 h, respectively, in liquid media, then plated in M3434 cytokine-enriched methylcellulose and maintained at the same concentration of doxocycline and 4OHT until the end of the experiment.

**qPCR.** RNA extraction was performed using a QIAGEN RNeasy Mini Kit. Complementary DNA (cDNA) was synthesized from RNA with a Superscript II kit (Invitrogen). Transcripts were amplified with Sybr Green PCR Master Mix (ABI). qPCR was performed on an ABI Prism real-time PCR system. The following primers were used to quantify *MLL4* expression: *MLL4*-F 5'-GCCACCTCTTGCGCTGT TCA-3'; *MLL4*-R 5'-ACACAACGCCAGCCCTTCAG-3'. The following primers were used to quantify *Prdx1*, *Cstb*, *Txnip* and *p21* expression: *Prdx1*-F 5'-GCGC TTCTGTGGATTCTCACTTCT-3'; *Prdx1*-R 5'-ACTCATATACTCTGAGCAA TGGTG-3'; *Cstb*-F 5'-GAAGTCCCAGCTTGAATCGAAAGAA-3'; *Cstb*-R 5'-TAGGAAGACAGGGTCAAAGGCTTGT-3'; *Txnip*-F 5'-GCTGCAACATCCT CAAAGTCGAA-3'; *Txnip*-R 5'-TCTTGAGAGTCGTCCACATCGTC-3'; *p21*-F 5'-CTGGGAGGGGACAAGAG-3' and *p21*-R 5'-GCTTGAGTGATAGAAAT CTG-3'. The *MLL4* flox and Cre-deleted alleles were quantified using the following primers: *MLL4*DNA-A 5'-AGGAACCTGAGGGAAACGAACC-3'; *MLL4*DNA-B 5'-GGAGAACAGGAGATGCCTCAGC-3'; *MLL4*DNA-C 5'-TGCAGAAGCC TGCTATGTCCAG-3'.

**shRNA targeting *MLL4* expression.** TransOMIC shRNAir against *MLL4* (RLGM-GU42557, target sequence: TGGGAATGATCTCTAAATGTT) and non-targeting control shRNA-mir (TRM1103, target sequence: ACCGGCTGAAGAGCCTGA TCA) were cloned from pMLP to the pLEPG backbone (a gift from J. Zuber). Retroviruses were generated in Phoenix-eco cells. WT MLL-AF9 cells were infected and selected for 96 h in puromycin (4  $\mu$ g ml<sup>-1</sup>), and depletion of *MLL4* messenger RNA (mRNA) was measured by qPCR using the primers for *MLL4* mRNA described above.

**RNA-seq.** To perform RNA-seq in MLL-AF9-infected cells, RNA extraction was performed using TRIzol (Ambion) following the manufacturer's protocol. RNA was washed, purified with an RNeasy kit (QIAGEN) and measured for quality using Agilent RNA 6000, Nano reagents and Bioanalyzer. RNA was then prepared



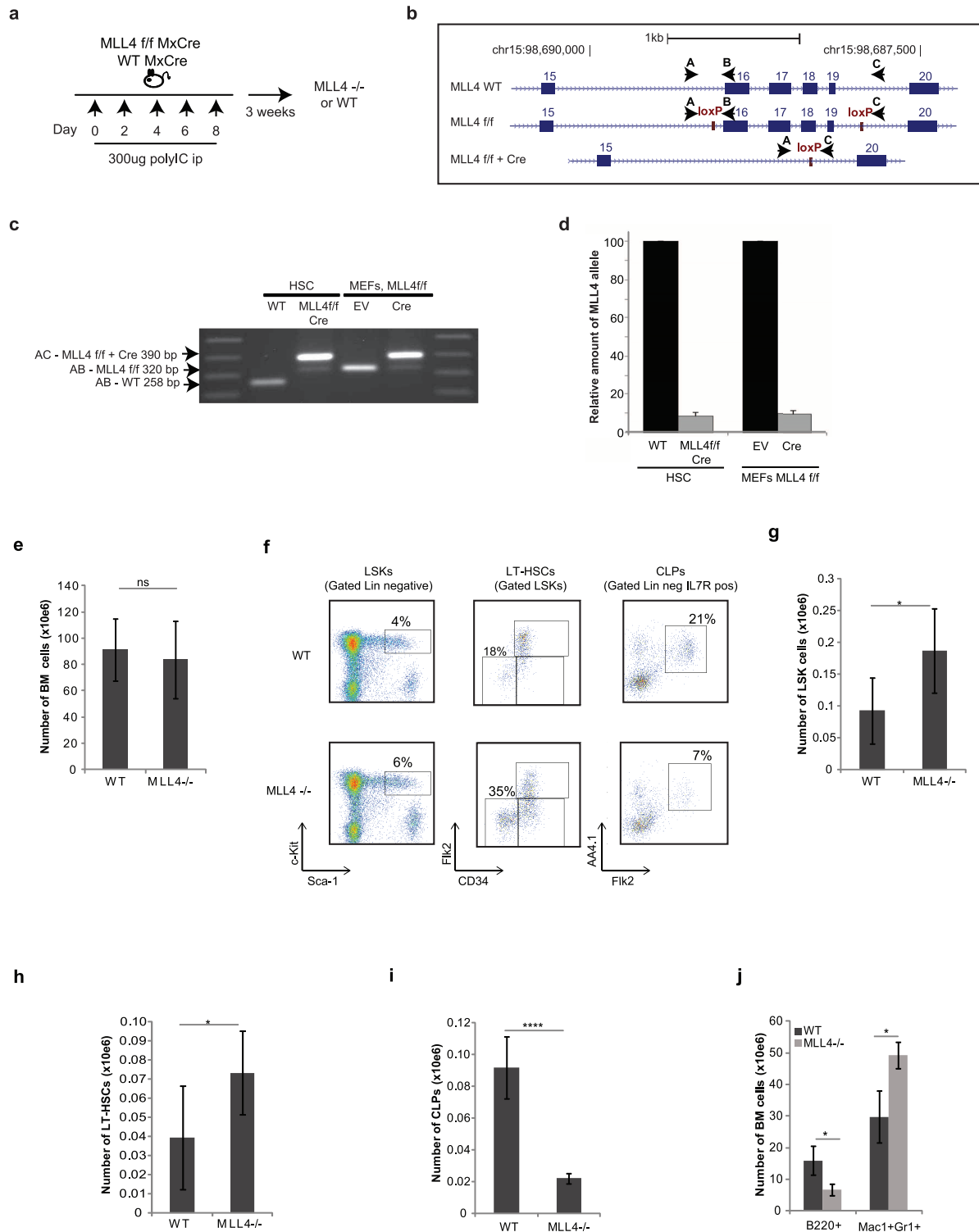
for sequencing using a TruSeq RNA sample prep kit (Illumina). To perform RNA-seq in LSK cells, we followed a protocol for single cell RNA-seq<sup>33</sup>. Sequence reads from each cDNA library were mapped onto the Build 37 assembly of the National Center for Biotechnology Information mouse genome data (July 2007; NCBI37/mm9) using TopHat, and outputted to bam format<sup>34</sup>. Bioconductor<sup>35</sup> packages were used to quantify the expression abundance of RefSeq genes from the aligned reads and calculate the reads per million on the genes' exons. For comparison of RNA-seq experiments in CRE-infected *MLL4<sup>fl/fl</sup>* cells transformed with MLL-AF9 versus CRE-infected WT cells transformed with MLL-AF9, genes with more than a two-fold change and FDR < 0.1 were designated as MLL4-dependent genes. Visualization was achieved by generating custom tracks for the University of California at Santa Cruz Genome Browser.

**Gene set analysis.** Gene sets were obtained from the biological processes classification from Gene Ontology<sup>36</sup>, Reactome<sup>37</sup> and Pathway Commons<sup>38</sup> to assess the canonical pathways and biological processes over-represented among the genes downregulated in *MLL4<sup>fl/fl</sup> MxCre* (*MLL4<sup>-/-</sup>*) versus WT *MxCre* (WT) LSK cells. *MLL-AF9* direct target genes were identified using the MLL-AF9 ChIP-seq data set deposited at the GEO under accession code GSE29130 and compared with MLL4-dependent genes in cells transformed with MLL-AF9. Fisher's exact test with Benjamini-Hochberg multiple-testing correction<sup>39</sup> was used as the measure of gene set over-representation. GSEA used Pre-ranked tool version 2.0.12 with 10,000 permutations. Moderated logarithmic fold change was used as the gene dysregulation ranking metric. MLL-AF9 'leukemic stem cell' self-renewal associated signatures (Up LSC and Down LSC) were defined on the basis of the microarray data set from ref. 16 deposited in the NCBI Gene Expression Omnibus under accession code GSE3725. FOXO1/3/4 positively regulated HSC gene sets were defined on the basis of analysing the microarray gene expression profiling of FOXO1/3/4-deficient HSC cells from the GEO data set GSE6623 (ref. 4). Limma package in Bioconductor<sup>40</sup> was used for microarray processing and differential gene expression. The glutathione-mediated detoxification pathway is a modified version of the genes annotated for the process at Pathway Commons<sup>38</sup>. The heatmap plot in Fig. 3j was generated on the genes in the glutathione, ROS and FOXO pathways that were associated with MLL4 deficiency identified by the leading edge of GSEA analyses unless indicated otherwise in the plot.

**Statistical analyses.** Multiple independent biological experiments were performed to assess the reproducibility of experimental findings. Each group is presented by mean  $\pm$  s.d. To compare two experimental groups, statistical tests were conducted using R statistical language (<http://r-project.org>). Unpaired, one-tailed *t*-tests were used for all analyses. For *ex vivo* experiments, multiple independent biological replicates were used. For leukaemia transplantation, five to seven recipients per group were used since variation among experiments was low. Animals were placed in different experimental groups and disease development was accessed blindly without prior knowledge of genotype. The significance between the longevity of cohorts was assessed by Kaplan-Meier survival analysis and log-rank (Mantel-Cox) tests. *P* values less than 0.05 were considered significant to reject the null hypothesis. No randomization was used in any experiment.

29. Celeste, A. *et al.* Histone H2AX phosphorylation is dispensable for the initial recognition of DNA breaks. *Nature Cell Biol.* **5**, 675-679 (2003).
30. Callen, E. *et al.* ATM prevents the persistence and propagation of chromosome breaks in lymphocytes. *Cell* **130**, 63-75 (2007).
31. Chiang, M. Y. *et al.* Leukemia-associated NOTCH1 alleles are weak tumor initiators but accelerate K-ras-initiated leukemia. *J. Clin. Invest.* **118**, 3181-3194 (2008).
32. Toledo, L. I. *et al.* A cell-based screen identifies ATR inhibitors with synthetic lethal properties for cancer-associated mutations. *Nature Struct. Mol. Biol.* **18**, 721-727 (2011).
33. Tang, F. *et al.* RNA-seq analysis to capture the transcriptome landscape of a single cell. *Nature Protocols* **5**, 516-535 (2010).
34. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-seq. *Bioinformatics* **25**, 1105-1111 (2009).
35. Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5**, R80 (2004).
36. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genet.* **25**, 25-29 (2000).
37. Vastrik, I. *et al.* Reactome: a knowledge base of biologic pathways and processes. *Genome Biol.* **8**, R39 (2007).
38. Cerami, E. G. *et al.* Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Res.* **39**, D685-D690 (2011).
39. Benjamini, Y. & Hochberg, Y. Controlling the False discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289-300 (1995).
40. Smyth, G. K. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Molec. Biol.* **3**, 3 (2004).

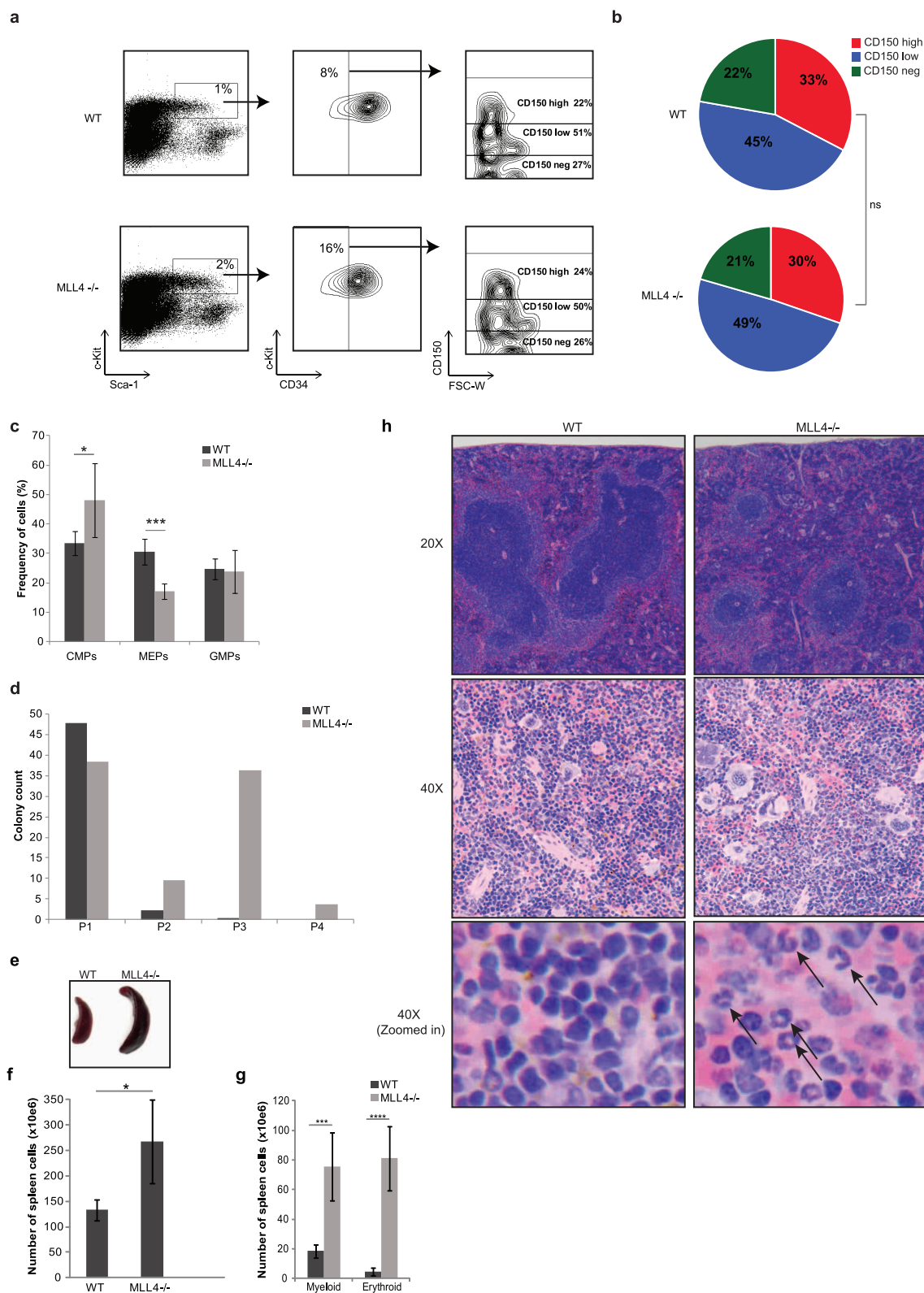




### Extended Data Figure 1 | Deletion of *MLL4* in haematopoietic stem cells.

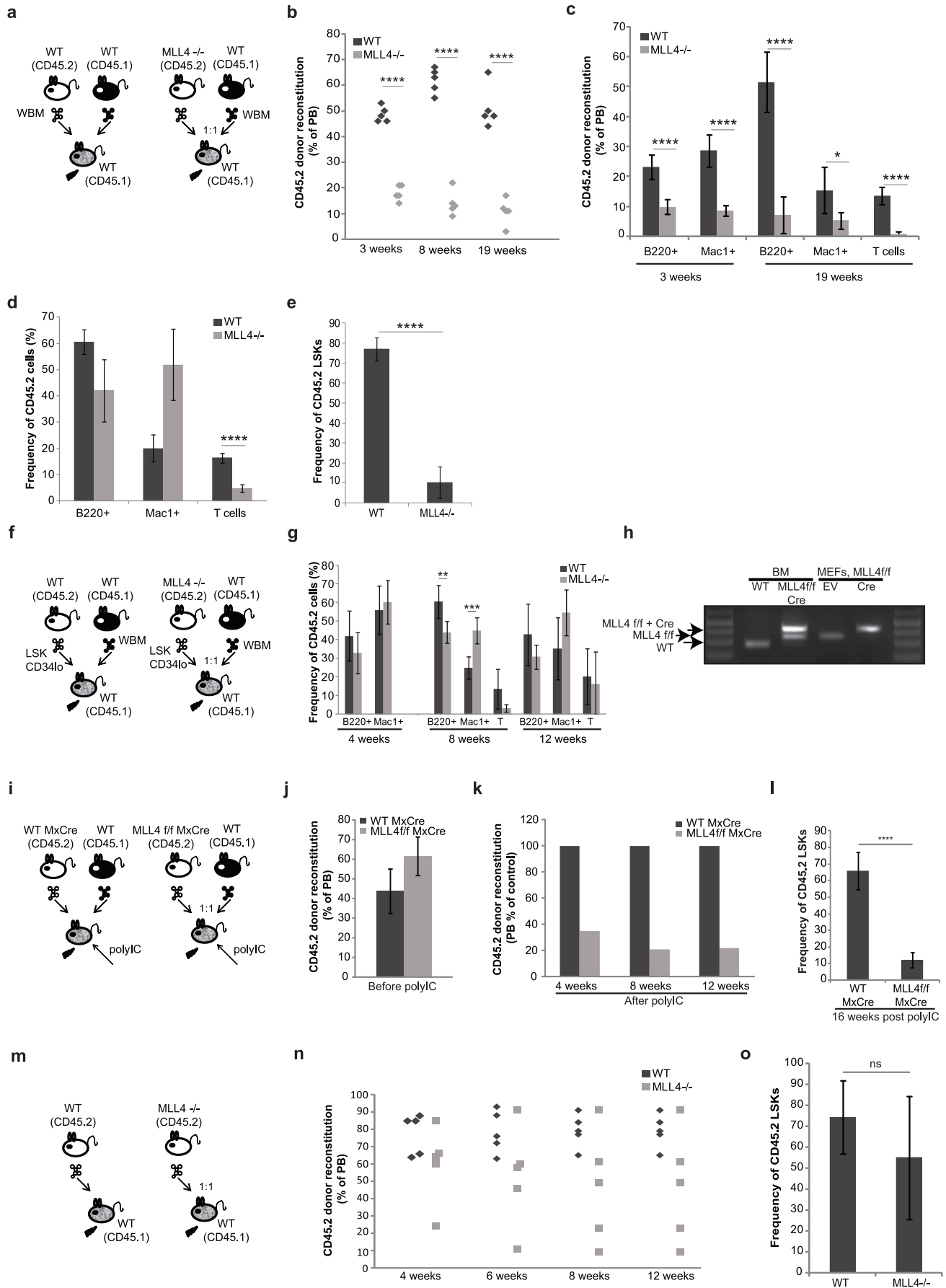
**a**, *MLL4*<sup>f/f</sup> mice (see methods) were crossed with the interferon-inducible transgene *MxCre* to obtain *MLL4*<sup>f/f</sup>*MxCre* and *WT**MxCre* mice. Animals were then treated with intraperitoneal (ip) injections of 300 µg of polyIC five times every other day and analysed 3 weeks after the last polyIC injection (generating WT and *MLL4*<sup>-/-</sup> mice). **b**, The *MLL4* wild-type locus (WT) and the floxed exons locus before and after Cre excision. Exons are represented in numbered boxes. The loxP sites (red rectangles) and the *MLL4* PCR primers (black arrows) are indicated. **c**, PCR analysis for conditional *MLL4* knockout mice (exons 16, 17, 18, 19). Genomic DNA from sorted HSCs derived from WT and *MLL4*<sup>-/-</sup> mice (PolyIC treated as in **a**) and from mouse embryonic fibroblasts (MEFs) derived from *MLL4*<sup>f/f</sup> cells infected with a retrovirus

expressing Cre or empty vector (EV) were analysed by PCR. The wild-type (258 base pairs (bp)) and floxed band (320 bp) were amplified with primers A and B, and the knockout band (390 bp) was amplified with primers A and C in different reactions. One of two independent genotyping experiments is shown. **d**, qPCR quantification of deletion efficiency in conditional knockout Cre-expressing cells. **e**, Whole bone-marrow cellularity 3 weeks after polyIC treatment of *WT**MxCre* and *MLL4*<sup>f/f</sup>*MxCre* mice (referred to as WT and *MLL4*<sup>-/-</sup> respectively). **f**, Representative FACS profiles pre-gated on live cells showing LSKs, LT-HSCs and common lymphoid progenitors, and quantification of these bone-marrow populations (**g-i**) as well as B cells and myeloid cells (**j**) in the bone marrow. All bar graphs show mean ± s.d. of at least three independent experiments.



**Extended Data Figure 2 | Extramedullary haematopoiesis in the absence of MLL4.** **a**, Representative FACS plot of LSKs separated based on CD34 and subsequently analysed by cell surface expression of CD150 (Slamf1). **b**, Pie charts summarize data from three independent mice of each genotype ( $\chi^2$  test,  $P > 0.8$ ). **c**, Frequency of cells determined by FACS analysis of Lin<sup>-</sup> Sca1<sup>+</sup> c-Kit<sup>-</sup> separated based on CD34 and Fc $\gamma$ RII/III. CMPs, common myeloid progenitors; MEPs, megakaryocyte-erythroid progenitors; GMPs, granulocyte macrophage progenitors. Mean  $\pm$  s.d. of three mice per group is shown. **d**, Quantification of colony numbers generated by WT and MLL4<sup>-/-</sup> myeloid

progenitors (LSK) in serial colony forming assays; y axis, number of colonies; x axis, serial assay, primary to quaternary. **e–g**, Splenomegaly and increased numbers of myeloid and erythroid cells in the spleens of MLL4<sup>-/-</sup> mice 3 weeks after polyIC treatment. Image of spleen representative of more than three independent experiments. **h**, Haematoxylin and eosin staining of spleen (upper panel,  $\times 20$ ; middle panel,  $\times 40$ ; lower panel, magnified picture of the middle panel to visualize details). Black arrows show presence of erythrocytes in MLL4-deficient spleens. Images were acquired in one experiment.

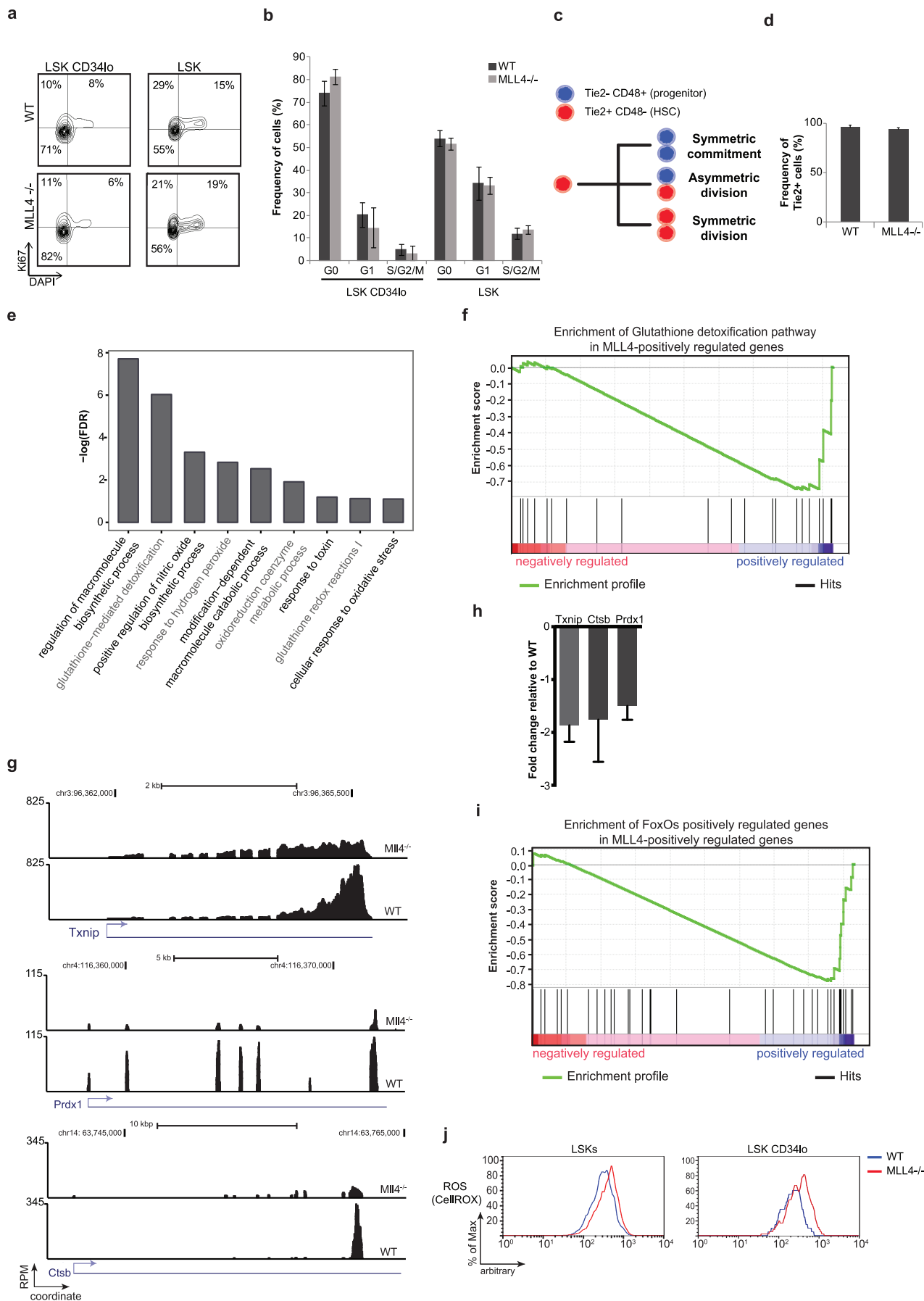


**Extended Data Figure 3 | MLL4-deficient HSCs have impaired**

**reconstitution capacity.** **a–e**, Whole bone-marrow (WBM) cells from WT or *MLL4*<sup>-/-</sup> mice (CD45.2) were mixed in 1:1 ratio with WT WBM (CD45.1) and transplanted into irradiated recipients (CD45.1). Reconstitution levels were monitored for 19 weeks after transplantation in the peripheral blood (PB) (**b**, **c**). **d**, Lineage distribution (B cells, myeloid-Mac1-positive cells and T cells) among remaining CD45.2 cells analysed in the peripheral blood 19 weeks after transplant. The total percentage of reconstitution as well as the frequency of the various lymphoid and myeloid subpopulations are severely diminished in the absence of MLL4 (**c**). However, among the few remaining *MLL4*<sup>-/-</sup> CD45.2-positive cells (**d**), there was a relatively higher frequency of myeloid cells and a diminished frequency of lymphoid cells. **e**, The frequency of donor-derived LSKs was determined in the bone marrow at week 19. Bar graphs show mean ± s.d. calculated from five mice of each genotype. **f**, Two hundred sort-purified LSK CD34<sup>lo</sup> cells (HSCs) from WT or *MLL4*<sup>-/-</sup> mice (CD45.2) were mixed with 500,000 WT WBM (CD45.1) and transplanted into irradiated recipients (CD45.1). **g**, Lineage distribution (B cells, myeloid-Mac1-positive cells and T cells) among remaining CD45.2 cells analysed in the peripheral blood at indicated time points. **h**, Genomic DNA derived from

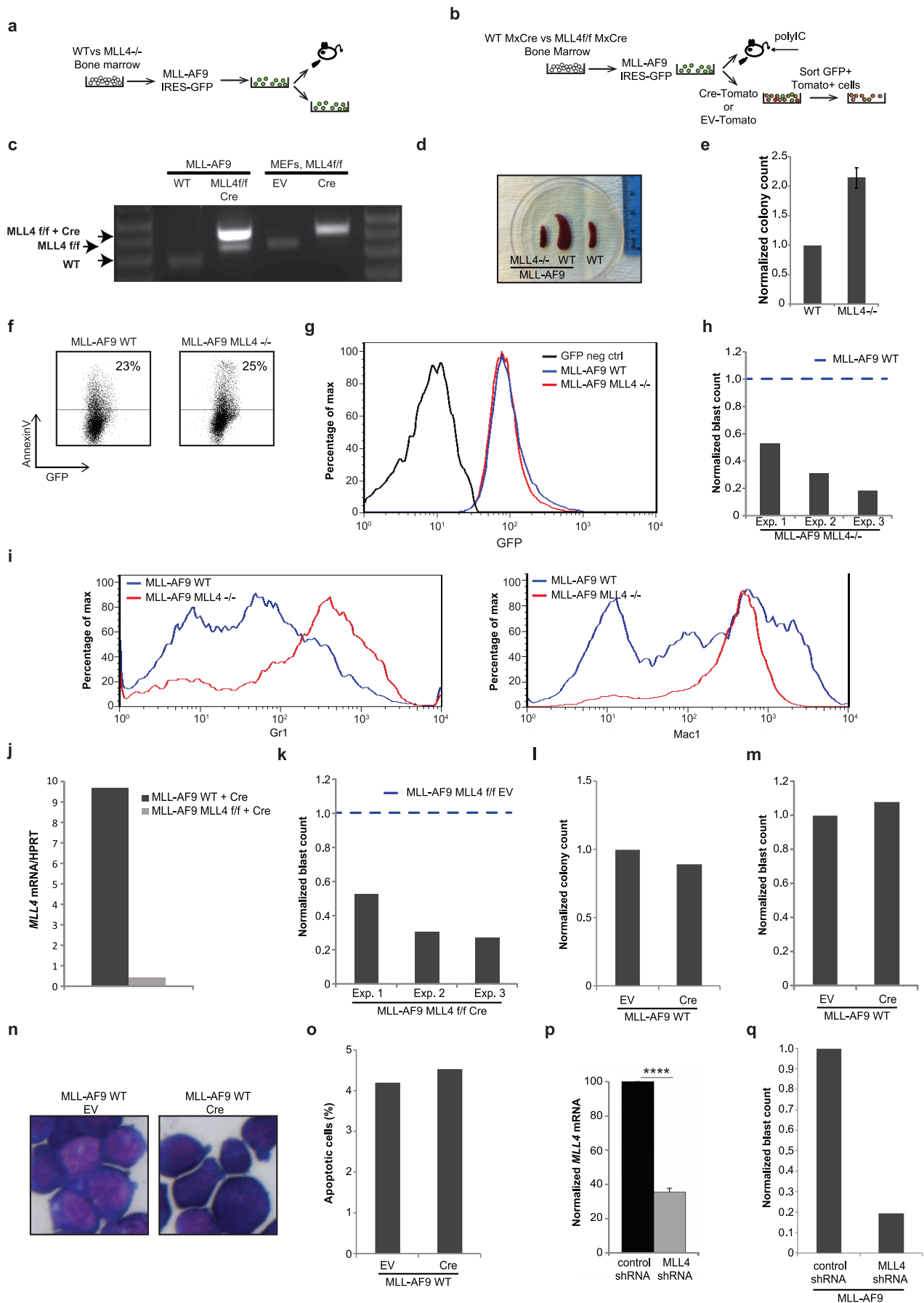
sorted bone marrow 14 weeks after transplant or from MEFs derived from *MLL4*<sup>f/f</sup> cells infected with a retrovirus expressing Cre or empty vector (EV) were analysed by PCR for *MLL4* deletion. Genotyping was performed once with CD45.2<sup>+</sup> cells pooled from the animals in each group at the end of the experiment (12 weeks). **i**, WBM from *WTMxCre* or *MLL4*<sup>f/f</sup>*MxCre* mice (CD45.2) were mixed in 1:1 ratio with WT WBM (CD45.1) and transplanted into irradiated recipient mice (CD45.1). **j**, Reconstitution levels were monitored in the peripheral blood at 5 weeks after transplantation and 1 day before the beginning of treatment with polyIC. **k**, Reconstitution levels were monitored in the peripheral blood at 4–12 weeks after polyIC treatment. **l**, The frequency of donor-derived LSKs was determined in the bone marrow at week 16. Bar graphs show mean ± s.d. calculated from five mice of each genotype. **m**, For non-competitive bone-marrow transplants, WBM cells (CD45.2) from WT or *MLL4*<sup>-/-</sup> mice (that is, WT *MxCre* and *MLL4*<sup>f/f</sup> *MxCre* mice 3 weeks after polyIC treatment) were transplanted into irradiated recipient mice (CD45.1). **n**, Reconstitution levels were monitored for 12 weeks after transplantation in the peripheral blood and the frequency of donor-derived LSKs (**o**) was determined in the bone marrow at week 16. Bar graphs show mean ± s.d. of five mice per group.





**Extended Data Figure 4 | MLL4 regulates the expression of genes in the glutathione- and FoxO-dependent pathways.** **a**, Representative FACs plots showing Ki-67 versus DAPI profiles in LSK CD34<sup>lo</sup> cells (left panel) and LSK cells (right panel). **b**, Summary of the cell-cycle profiles as in **a** for three independent mice per group. **c**, Schematic of division patterns of HSCs. **d**, Tie2 expression in bone-marrow CD150<sup>+</sup>CD48<sup>+</sup>CD41<sup>+</sup>Flt3<sup>+</sup>CD34<sup>+</sup> LSK cells from WT or *MLL4*<sup>-/-</sup> mice. **e**, Canonical pathways and biological processes over-represented within the 1,000 most downregulated genes in *MLL4*<sup>fl/fl</sup> MxCre (*MLL4*<sup>-/-</sup>) relative to WT MxCre (WT) sorted LSKs. **f**, GSEA shows enrichment of glutathione detoxification pathway in WT MxCre (WT) relative

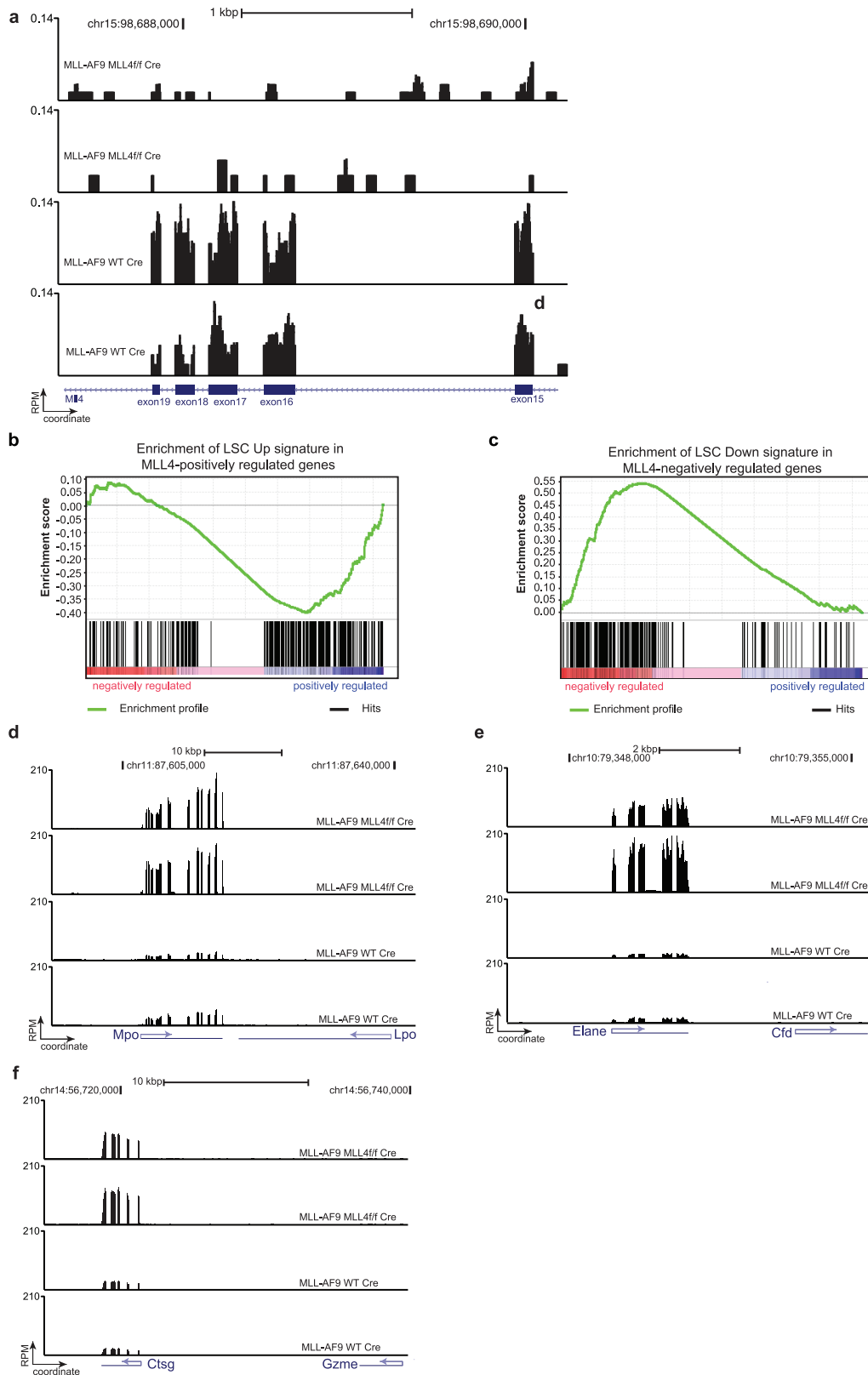
to *MLL4*<sup>fl/fl</sup> MxCre (*MLL4*<sup>-/-</sup>) LSKs (FDR < 0.1). **g**, RNA-seq read histograms at *Txnip*, *Prdx1* and *Ctsb*. The *x* axis represents the linear sequence of genomic DNA; the *y* axis represents the reads per million aligned reads (RPM). The genomic scale in kilobase pairs (kbp) is indicated above the tracks. **h**, mRNA levels detected by qRT-PCR in purified HSCs of selected genes (*Txnip*, *Prdx1* and *Ctsb*) that were downregulated in the absence of MLL4. **i**, GSEA plot shows downregulation of the FOXO1/3/4 positively regulated genes in *MLL4*<sup>fl/fl</sup> MxCre (*MLL4*<sup>-/-</sup>) LSKs (FDR < 0.1). **j**, LSK and LSK CD34<sup>lo</sup> cells from WT or *MLL4*<sup>-/-</sup> mice were stained with CellROX Deep Red Reagent to measure the levels of ROS. One representative of three experiments is shown.



**Extended Data Figure 5 | MLL4 is required for MLL–AF9 transformation *in vivo* and *in vitro*.** **a**, WT and *MLL4*<sup>−/−</sup> bone-marrow cells were transformed with MLL–AF9 and injected into irradiated recipients (650 rad) or maintained in culture for *in vitro* experiments. **b**, Non-polyIC-treated *WT*MxCre and *MLL4*<sup>f/f</sup> MxCre bone-marrow cells were transformed with MLL–AF9. Cells were subsequently infected with retrovirus containing Cre-recombinase (CRE-Tomato) or injected into mice that were administered polyIC 1 week later. **c**, PCR analysis of genomic DNA shows the extent of *MLL4* deletion in MLL–AF9-infected cells. *MLL4*<sup>f/f</sup> and *MLL4*<sup>f/f</sup> Cre-infected MEFs were used as a control. Genotyping was performed once. **d**, Spleens from mice 29 days after injection with WT MLL–AF9 or *MLL4*<sup>−/−</sup> MLL–AF9 cells, and spleen from non-injected littermates (WT) (see also Fig. 2a–c). Photographs were taken in one experiment. **e**, Normalized colony counts scored 11 days after culture of WT or *MLL4*<sup>−/−</sup> whole bone marrow (non-transformed) in semi-solid media in the presence of IL3, IL6 and SCF. **f**, Representative FACS plots showing AnnexinV versus GFP staining in MLL–AF9 WT or *MLL4*<sup>−/−</sup> cells cultured in semi-solid media (as in **a**). **g**, Histogram of GFP expression 10 days after MLL–AF9 transformation. **h**, Frequency of cells identified as blasts evaluated from cytopsin samples in Fig. 2d. Data are shown normalized to WT counts (dotted line) in three

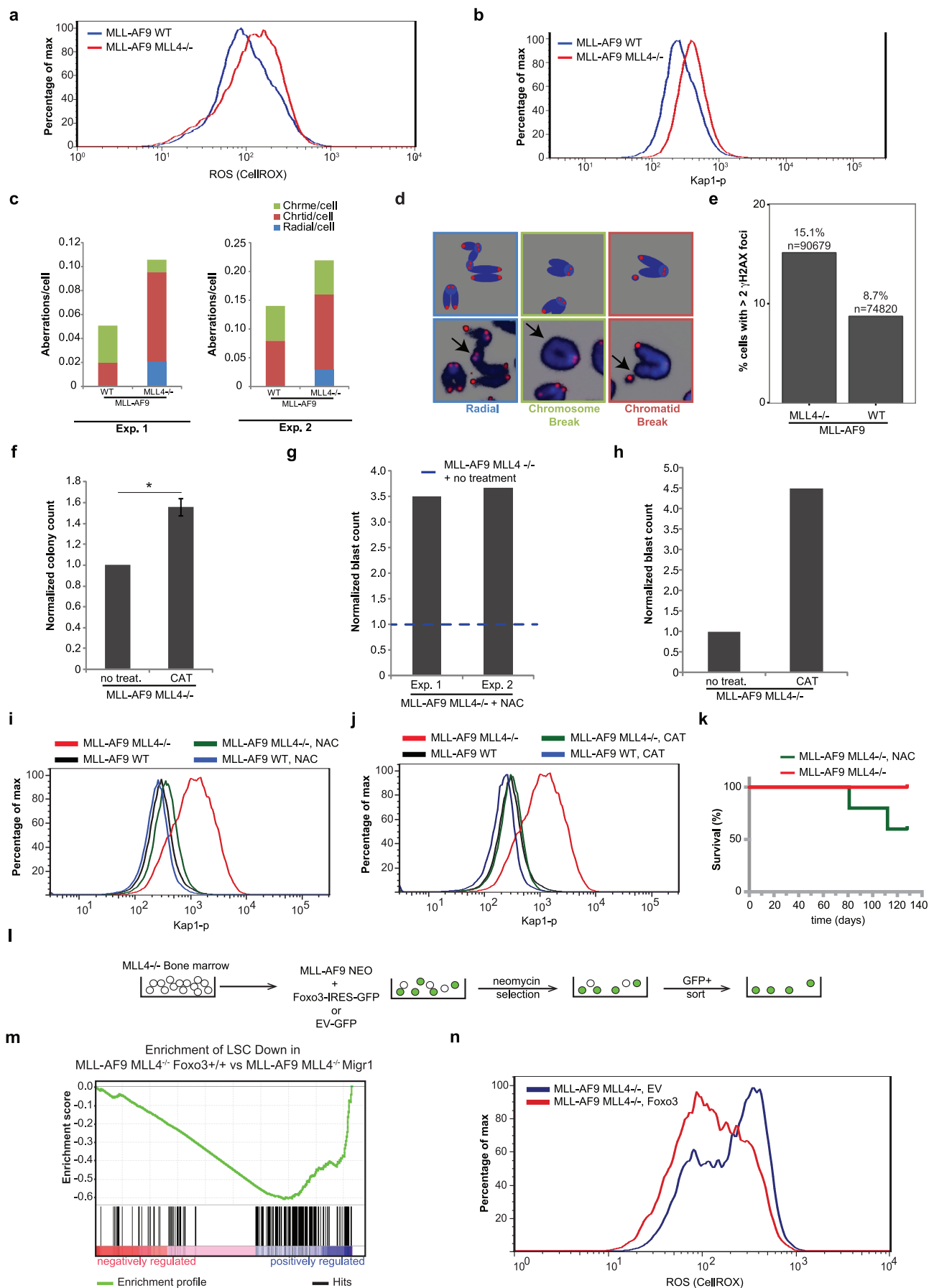
independent experiments. **i**, The morphological changes observed in *MLL4*-deficient MLL–AF9 cells are accompanied by increased expression of the myeloid markers Mac1 (right) and Gr1 (left). **j**, WT or *MLL4*<sup>f/f</sup> bone-marrow cells were transformed with MLL–AF9 and subsequently *MLL4* was excised by retroviral expression of CRE as in **b**. Five days later, *MLL4* mRNA levels were measured by qPCR. **k**, Frequency of cells identified as blasts in the cytopsin samples in Fig. 2h. Data are shown normalized to WT counts (dotted line) in three independent experiments. **l–o**, WT bone-marrow cells were transformed with MLL–AF9 and subsequently infected with retroviruses expressing CRE (as in **b**). Blast colony counts (**l**), frequency of blasts evaluated by May–Grünwald–Giemsa stained cytopsin (**m**, **n**) and frequency of apoptotic cells determined by morphology (**o**) were calculated after culture in semi-solid media supplemented with SCF, IL3 and IL6. Images of cytopsin were acquired once. **p**, WT MLL–AF9 cells were stably infected with a retrovirus encoding an shRNA to target and silence *MLL4* expression. shRNA depletion of *MLL4* mRNA was measured by qPCR and normalized to levels in non-target control shRNA-infected cells. **q**, Frequency of cells identified as blasts in the cytopsin samples from *MLL4* shRNA-infected cells compared with control shRNA-infected cells.





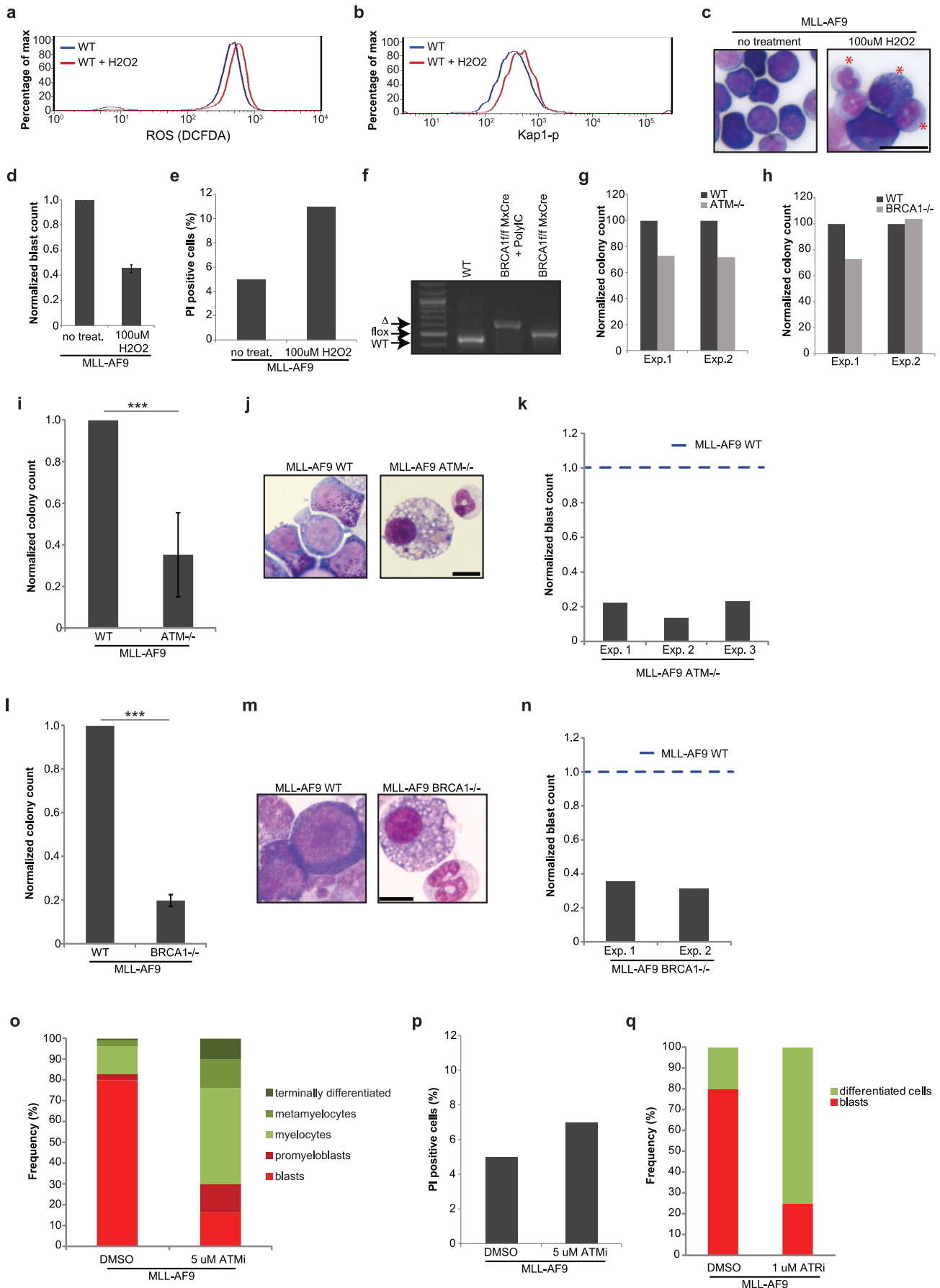
**Extended Data Figure 6 | Genes associated with myeloid maturation are significantly upregulated after *MLL4* deletion in cells transformed with MLL-AF9. a**, RNA-seq read histograms at the *MLL4* exons 16–19 in MLL-AF9 *MLL4*<sup>fl/fl</sup> Cre and MLL-AF9 WT Cre cells. **b**, LSC Up gene set constitutes genes upregulated in the MLL-AF9 ‘leukemic stem cell’ self-renewal associated signature. GSEA plot demonstrates downregulation of LSC Up gene set in MLL-AF9 *MLL4*<sup>fl/fl</sup> Cre cells (FDR < 0.1). **c**, LSC Down gene set constitutes genes downregulated in the MLL-AF9 ‘leukemic stem cell’ self-renewal

associated signature. GSEA plot demonstrates upregulation of LSC Down gene set in MLL-AF9 *MLL4*<sup>fl/fl</sup> Cre cells (FDR < 0.1). **d–f**, Comparison of RNA-seq read histograms at the genes *Mpo* (myeloperoxidase), *Elane/Elas2* (neutrophil elastase) and *Ctsg* (cathepsin G) in MLL-AF9 *MLL4*<sup>fl/fl</sup> Cre and MLL-AF9 WT Cre cells. The x axis represents the linear sequence of genomic DNA; the y axis represents the reads per million aligned reads (RPM). The genomic scale in kilobase pairs (kbp) is indicated above the tracks.



**Extended Data Figure 7 | Increased levels of DNA damage and ROS in *MLL4*-deficient cells transformed with *MLL*-AF9.** **a**, WT and *MLL4*<sup>-/-</sup> cells were stained with CellROX Deep Red Reagent to measure the levels of ROS after *MLL*-AF9 infection. **b**, The levels of phosphorylated Kap1 were determined by flow cytometry. One representative of at least three independent measurements is shown. **c**, Levels of aberrations (chromosome breaks, chromatid breaks and radial chromosomes) in metaphase spreads in two independent experiments derived from *MLL*-AF9 WT and *MLL4*<sup>-/-</sup> cells. **d**, Examples of chromosome aberrations. One representative of two experiments. **e**, High-throughput microscopy imaging of *MLL4*<sup>-/-</sup> (*n* = 90,679) and WT (*n* = 74,820) quantifies the percentage of cells with at least three  $\gamma$ -H2AX foci. On average, 1.3 and 5.9 foci per cell were observed in WT and *MLL4*<sup>-/-</sup> *MLL*-AF9 cells, respectively. **f–j**, After infection with *MLL*-AF9, *MLL4*-deficient cells were expanded in semi-solid media in the presence or absence of the antioxidants NAC or catalase. **f**, Normalized colony counts with or without catalase treatment. Data show mean  $\pm$  s.d. of three independent experiments. **g**, Frequency of cells identified as blasts with or without NAC in the cytopspins of Fig. 3e. Data are shown normalized to WT counts (dotted line) in two independent experiments. **h**, Frequency of blasts with or without catalase treatment were quantified on the basis of morphology. **i, j**, NAC or catalase treatment reduces the levels of phosphorylated Kap1 in

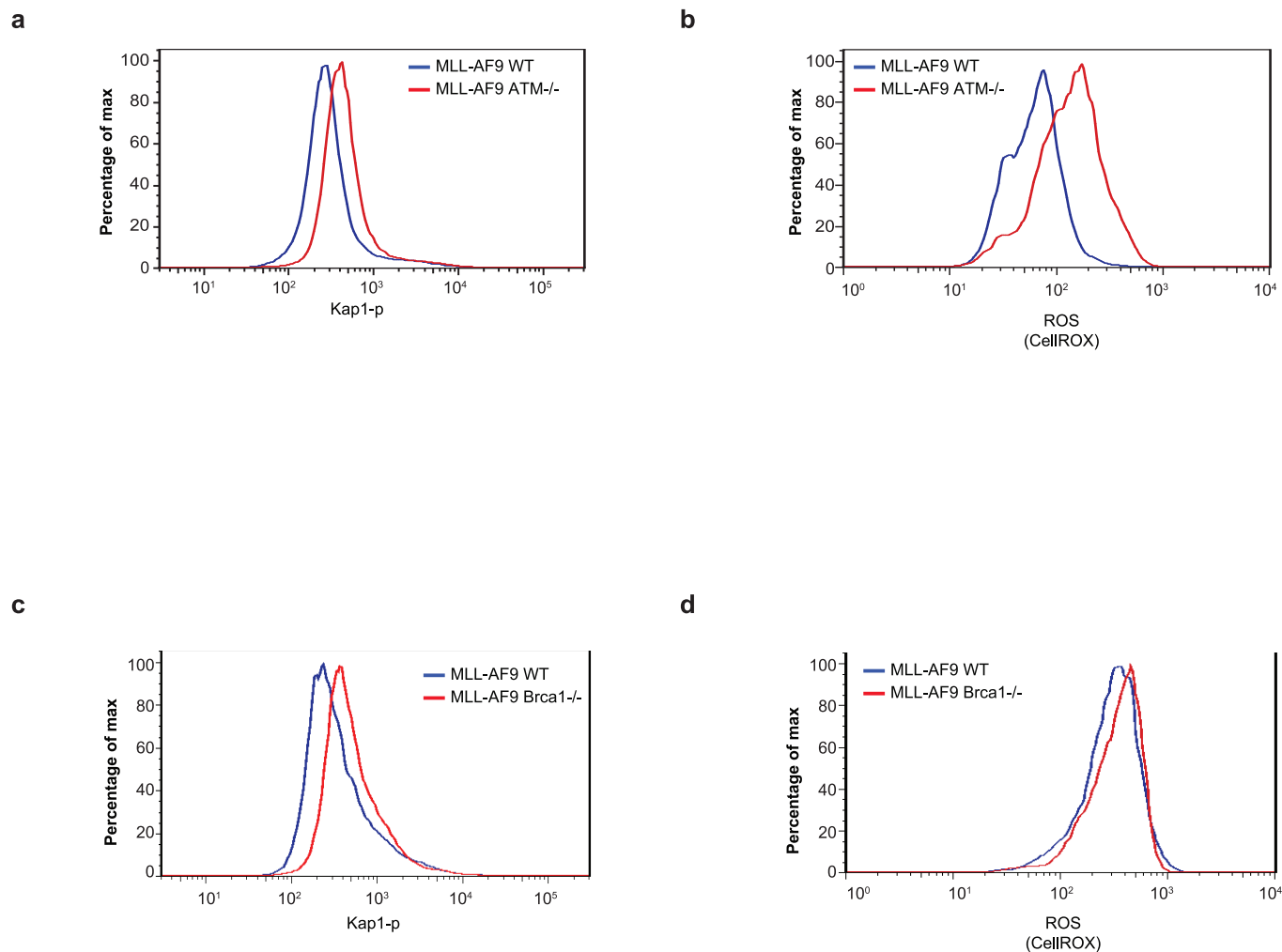
*MLL4*-deficient *MLL*-AF9 cells. The two different treatments (NAC, catalase) and controls (red and black lines) were performed in the same experiment but controls are plotted separately in **i** and **j** for simplicity. **k**, Bone-marrow cells from *MLL4*<sup>fl/fl</sup> *MxCre* mice (without polyIC treatment) were collected 4 days after 5-FU treatment and infected with a retrovirus containing *MLL*-AF9. After expansion in semi-solid media, cells were injected into mice that were subsequently (1 week later) administered polyIC to excise *MLL4* *in vivo*. One group of mice was fed with NAC in the drinking water starting 1 week before the injection of the transformed cells. Two out of five animals treated with NAC died (both animals displayed elevated white blood cell counts at time of death) and none of the untreated mice died (and all had normal white blood cell counts). Survival curves were determined at the indicated time points; *n* = 5 mice per group. **l–n**, *MLL4*<sup>-/-</sup> bone-marrow cells were co-transformed with *MLL*-AF9-neo and either empty vector or retroviruses encoding FOXO3-IRES-GFP. One week after selection in G418, GFP<sup>+</sup> cells were sorted, then cultured *ex vivo*. **m**, GSEA plot demonstrates that FOXO complementation reversed upregulation of LSC Down gene set in *MLL4*<sup>fl/fl</sup>-Cre *MLL*-AF9 cells (FDR < 0.1). **n**, Cells were stained with CellROX Deep Red Reagent to measure the levels of ROS. One representative of two independent measurements is shown.





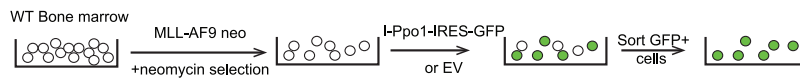
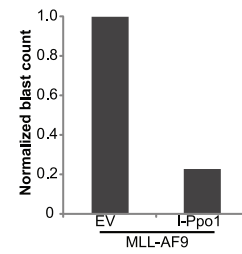
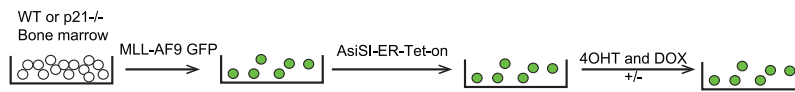
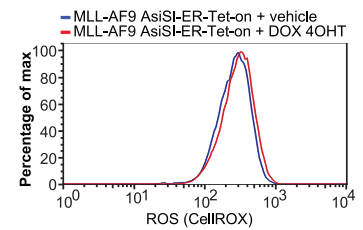
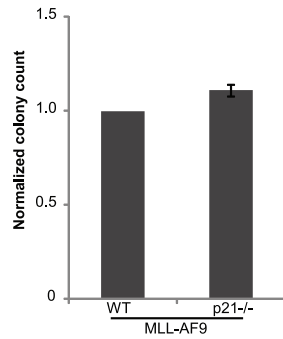
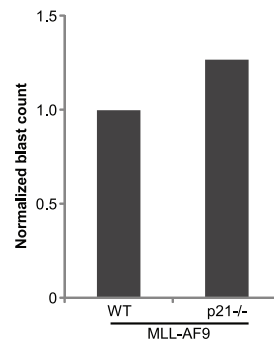
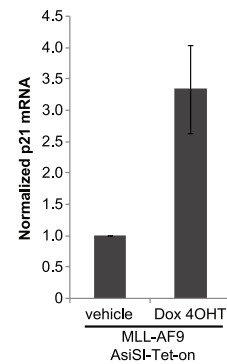
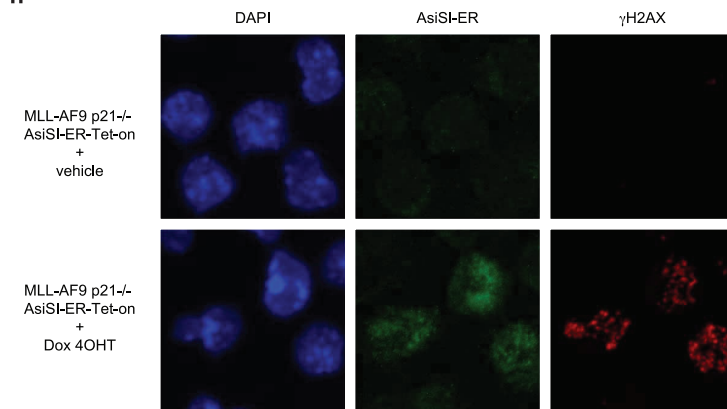
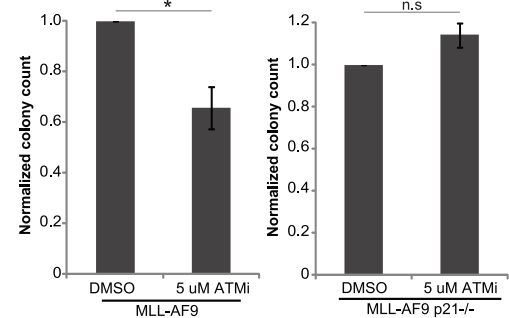
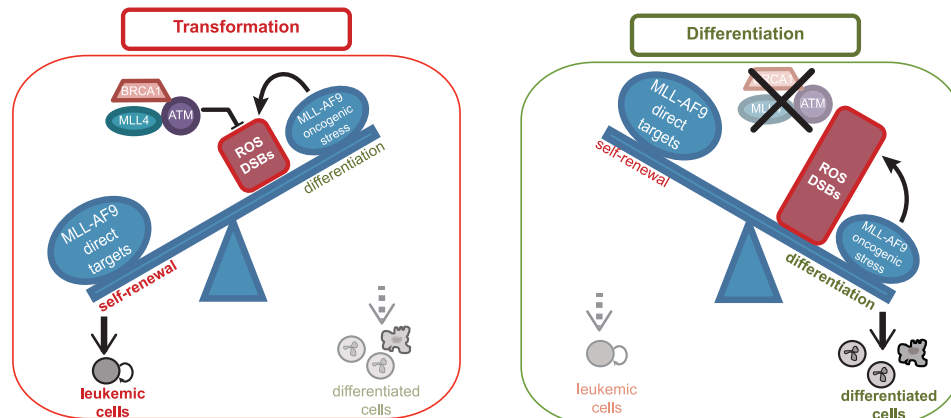
**Extended Data Figure 8 | H<sub>2</sub>O<sub>2</sub> treatment, ATM-, BRCA1-deficiency or ATM and ATR inhibition leads to myeloid differentiation of cells transformed with MLL-AF9 *ex vivo*.** **a–e**, WT cells transformed with MLL-AF9 (Extended Data Fig. 5a) and expanded in semi-solid media. Cells were treated with 100  $\mu$ M H<sub>2</sub>O<sub>2</sub> and the levels of ROS detected by DCF-DA staining (**a**) and phosphorylated Kap1 (**b**) were determined 48 h after treatment. One representative of two independent experiments is shown. The number of cells with blast morphology (**c**) was quantified 48 h after treatment (**d**). Red stars in (**c**) indicate cells with morphological changes characteristic of differentiation; images of one out of two independent experiments. The same WT controls were used in Fig. 2d. Bar graph shows mean  $\pm$  s.d. of two independent measurements. **e**, Bar graph shows the frequency of propidium iodide (PI)-positive cells 48 h after 100  $\mu$ M H<sub>2</sub>O<sub>2</sub> treatment. **f**, Genomic DNA was extracted from WT and *BRCA1*<sup>f/f</sup> *MxCre* bone-marrow cells from mice treated or not with polyIC. Detection of the WT-, deleted- and floxed alleles of *BRCA1* are indicated. **g, h**, Whole bone-marrow cells of the indicated genotypes were cultured in methylcellulose media supplemented with IL3, IL6 and SCF and the colony numbers were scored between days 8 and 10. WT or *ATM*<sup>-/-</sup> cells were transformed with MLL-AF9 as shown in Extended Data Fig. 5a, and blast colony counts (**i**) and morphology in methylcellulose,

May-Grünwald-Giemsa stains (**j**) were determined 10 days later; images of one out of three experiments. Bar graph shows mean  $\pm$  s.d. of four independent experiments. **k**, The frequency of cells identified as blasts in (**j**) was determined. Data are shown normalized to WT counts (dotted line) in three independent experiments. **l–n**, Three weeks after polyIC treatment of *WTMxCre* and *BRCA1*<sup>f/f</sup> *MxCre* mice, bone-marrow cells were transformed with MLL-AF9, and blast colony counts (**l**) and colony morphology (**m**) were assessed at day 10 after transformation (images of one out of two experiments). Bar graph shows mean  $\pm$  s.d. of three independent experiments. **n**, The frequency of cells identified as blasts is shown normalized to WT counts (dotted line) in two independent experiments. **o**, WT cells transformed with MLL-AF9 (as in Extended Data Fig. 5a) were expanded in semi-solid media, treated with 5  $\mu$ M of ATMi or vehicle for 48 h, and the frequency of blasts and cells at different stages of differentiation was determined by morphology. **p**, Cells transformed with MLL-AF9 were treated with 5  $\mu$ M of ATMi for 48 h, and the frequency of propidium iodide (PI)-positive cells is plotted. **q**, Cells transformed with MLL-AF9 were treated with 1  $\mu$ M of ATRi for 24–48 h and the frequency of blasts was determined. One of two representative experiments is shown. Scale bars, 10  $\mu$ m.



**Extended Data Figure 9 | ROS and DNA damage in *ATM*- and *BRCA1*-deficient cells transformed with MLL-AF9.** **a, b,** WT and *ATM*<sup>-/-</sup> cells were transformed with MLL-AF9 and the levels of phosphorylated Kap-1 (**a**) and ROS (detected by staining cells with CellROX Deep Red reagent) (**b**) were measured 10 days after expansion in semi-solid media. **c, d,** WT and *BRCA1*<sup>-/-</sup>

bone-marrow cells were transformed with MLL-AF9 and the levels of phosphorylated Kap1 (**c**) and ROS (**d**) were measured 10 days after expansion in semi-solid media. One representative of two independent experiments is shown.

**a****b****c****d****e****f****g****h****i****j**

**Extended Data Figure 10 | DNA damage induces p21-dependent**

**differentiation of cells transformed with MLL-AF9.** **a, b**, WT cells were transformed with MLL-AF9-neo. After 2 weeks of selection in neomycin-supplemented media, cells were infected with retroviruses expressing I-Ppo1 or empty vector. **b**, Eight to ten days after sorting GFP<sup>+</sup> cells, the frequency of I-Ppo1 infected cells identified as blasts in the cytopins of Fig. 4d was determined and normalized to the frequency of blasts in empty vector infected cells. **c, d**, WT cells were transformed with MLL-AF9 and then were infected with a retrovirus containing the inducible restriction enzyme AsiSi-ER-Tet-on. After selection, cells were treated with 4OHT and doxocycline for 24 h and the levels of ROS were measured by CellROX staining. **e, f**, WT or *p21*<sup>-/-</sup> cells were transformed with MLL-AF9. Colony counts (**e**) and

frequency of blasts (morphology in May-Grünwald-Giemsa stains) (**f**) were determined 8–10 days later. **g, h**, WT or *p21*<sup>-/-</sup> cells were transformed with MLL-AF9 and then infected with a retrovirus containing the restriction enzyme AsiSi. **g**, After selection, cells were treated with 4OHT and doxocycline for 24 h and the levels of *p21* mRNA were measured in WT cells by RT-qPCR. **h**, *p21*<sup>-/-</sup> cells containing AsiSi were treated with 4OHT and doxocycline for 24 h, and  $\gamma$ -H2AX foci (red) and AsiSi-ER staining (green) were examined by immunofluorescence (images of one out of two experiments). **i**, The number of colonies in WT and *p21*<sup>-/-</sup> MLL-AF9 cells 5–7 days after treatment with ATMi. **j**, Model showing that genome caretakers MLL4, ATM and BRCA1 prevent differentiation by protecting against ROS and DSBs.



# Endothelial-cell FAK targeting sensitizes tumours to DNA-damaging therapy

Bernardo Tavora<sup>1</sup>, Louise E. Reynolds<sup>1\*</sup>, Silvia Batista<sup>1\*</sup>, Fevzi Demircioglu<sup>1\*</sup>, Isabelle Fernandez<sup>1\*</sup>, Tanguy Lechertier<sup>1\*</sup>, Delphine M. Lees<sup>1\*</sup>, Ping-Pui Wong<sup>1\*</sup>, Annika Alexopoulou<sup>1</sup>, George Elia<sup>2</sup>, Andrew Clear<sup>3</sup>, Adeline Ledoux<sup>4</sup>, Jill Hunter<sup>4</sup>, Neil Perkins<sup>4</sup>, John G. Gribben<sup>3</sup> & Kairbaan M. Hodivala-Dilke<sup>1</sup>

**Chemoresistance is a serious limitation of cancer treatment<sup>1</sup>. Until recently, almost all the work done to study this limitation has been restricted to tumour cells<sup>2</sup>. Here we identify a novel molecular mechanism by which endothelial cells regulate chemosensitivity. We establish that specific targeting of focal adhesion kinase (FAK; also known as PTK2) in endothelial cells is sufficient to induce tumour-cell sensitization to DNA-damaging therapies and thus inhibit tumour growth in mice. The clinical relevance of this work is supported by our observations that low blood vessel FAK expression is associated with complete remission in human lymphoma. Our study shows that deletion of FAK in endothelial cells has no apparent effect on blood vessel function *per se*, but induces increased apoptosis and decreased proliferation within perivascular tumour-cell compartments of doxorubicin- and radiotherapy-treated mice. Mechanistically, we demonstrate that endothelial-cell FAK is required for DNA-damage-induced NF- $\kappa$ B activation *in vivo* and *in vitro*, and the production of cytokines from endothelial cells. Moreover, loss of endothelial-cell FAK reduces DNA-damage-induced cytokine production, thus enhancing chemosensitization of tumour cells to DNA-damaging therapies *in vitro* and *in vivo*. Overall, our data identify endothelial-cell FAK as a regulator of tumour chemosensitivity. Furthermore, we anticipate that this proof-of-principle data will be a starting point for the development of new possible strategies to regulate chemosensitization by targeting endothelial-cell FAK specifically.**

Despite encouraging initial responses to DNA-damaging chemotherapies and radiotherapy, many tumours become resistant to treatment<sup>1</sup>. Previous work has concentrated on understanding resistance by focusing on mechanisms within tumour cells<sup>2</sup>. Although recent evidence suggests that the tumour stroma can regulate chemoresistance, the underlying molecular mechanisms are largely unknown<sup>3–9</sup>.

FAK is a non-receptor tyrosine kinase and regulator of cell migration, proliferation and survival<sup>10</sup>. FAK can also regulate transcription via its scaffolding functions in the nucleus<sup>11</sup>. Although some studies have implicated a role for endothelial-cell FAK in tumour growth and angiogenesis<sup>12–14</sup>, its role in the regulation of chemoresistance has not been identified.

Here we demonstrate that targeting endothelial-cell FAK, in established tumours, is sufficient to sensitize tumour cells to DNA-damaging therapies. *Pdgfrb-iCre<sup>ER</sup>;Fak<sup>fl/fl</sup>* mice were injected subcutaneously with mouse melanoma (B16F0) or lung carcinoma (CMT19T) cell lines. At 7 days after tumour-cell inoculation, endothelial-cell FAK deletion was induced, generating ECFAK<sup>KO</sup> mice (Extended Data Fig. 1). Mice were then treated with one of two forms of DNA-damaging therapies: doxorubicin or radiation. Similarly treated *Pdgfrb-iCre<sup>ER</sup>*;non-floxed mice or *Fak<sup>fl/fl</sup>* mice (ECFAK<sup>WT</sup>) were used as controls for endothelial-cell FAK expression. Loss of endothelial-cell FAK did not affect B16F0 or CMT19T tumour growth in placebo-treated or non-irradiated mice

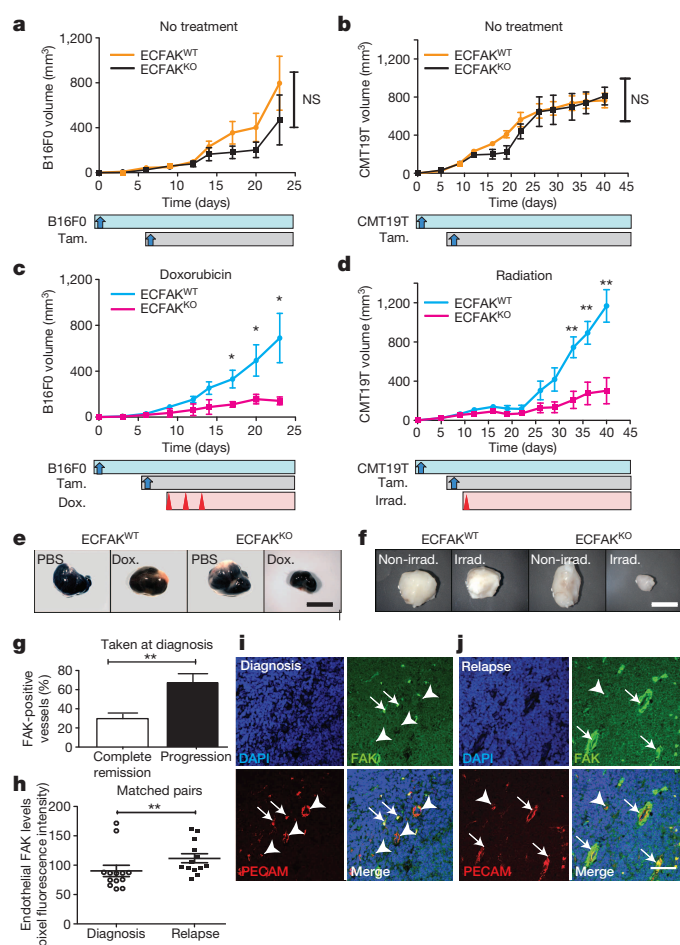
(Fig. 1a, b), nor did it affect tumour angiogenesis, blood vessel perfusion, or endothelial-cell apoptosis *in vivo* (Extended Data Fig. 2). In contrast to deleting endothelial-cell FAK before tumour development<sup>14</sup>, here our data indicate that endothelial-cell FAK deletion after tumour growth has begun is not sufficient to affect blood vessel density, results that are supported by other studies<sup>15,16</sup>. Moreover, we go on to show that doxorubicin or radiation therapy in ECFAK<sup>WT</sup> mice was not sufficient to affect B16F0 or CMT19T tumour growth, respectively, indicating that these tumour types are not sensitive to such forms of therapy *in vivo* (Fig. 1c, d). In contrast, endothelial-cell FAK deletion resulted in sensitizing B16F0 tumours to doxorubicin, causing a significant delay in tumour growth when compared with similarly treated ECFAK<sup>WT</sup> mice (Fig. 1c). Likewise, endothelial-cell FAK deletion in mice bearing CMT19T tumours sensitized tumours to radiation therapy, also leading to a significant decrease in tumour growth rates (Fig. 1d). Despite elevated numbers of  $\gamma$ H2AX-positive tumour-cell nuclei (an indicator of DNA damage) in ECFAK<sup>KO</sup> when compared with ECFAK<sup>WT</sup> mice after treatment (Extended Data Fig. 3a), no changes in tumour blood vessel permeability, doxorubicin delivery, tumour hypoxia or CD45-positive immune-cell infiltration were observed between genotypes (Extended Data Fig. 3b–e). These data suggest that loss of endothelial-cell FAK enhances tumour-cell responses to DNA damage without affecting the delivery function of blood vessels. Indeed, using other mouse models of cancer—experimental metastasis to the lung, using either tail-vein injection of B16F10 melanoma or EuMycBCL2 lymphoma—we show that loss of endothelial-cell FAK is sufficient to sensitize tumours to doxorubicin and significantly extend median survival (Extended Data Fig. 4). Together, these data demonstrate that endothelial-cell FAK deletion alone is sufficient to sensitize tumours to DNA-damaging therapies.

The clinical relevance of our results is also apparent in human cancer. Chemotherapy can induce complete remission and be curative in some subsets of lymphoma. However, many patients are resistant to doxorubicin-containing chemotherapy, and either fail to achieve remission or relapse and subsequently show progression months or even years later<sup>17,18</sup>. We asked whether disease progression in lymphoma patients correlated with altered blood-vessel FAK expression. First, sections of human lymphoma samples, all taken at diagnosis, were analysed for the percentage of FAK-positive blood vessels. Results showed that, after doxorubicin-containing treatment, a high percentage of FAK-positive blood vessels was associated significantly with subsequent disease progression, while a low percentage of FAK-positive blood vessels correlated significantly with complete remission (Fig. 1g). This result is unlikely to be due to a direct effect of doxorubicin treatment on FAK expression levels because, although doxorubicin can affect FAK localization in endothelial cells, suggesting a possible change in function, it does not affect FAK expression levels (Extended Data Fig. 5). Second, using

<sup>1</sup>Adhesion and Angiogenesis Laboratory, Centre for Tumour Biology, Barts Cancer Institute, CR-UK Centre of Excellence, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK.

<sup>2</sup>Barts Cancer Institute, CR-UK Centre of Excellence, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK. <sup>3</sup>Centre for Haemato-Oncology, Barts Cancer Institute, CR-UK Centre of Excellence, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK. <sup>4</sup>Institute for Cell and Molecular Biosciences (ICaMB), Medical School, Newcastle University, Catherine Cookson Building, Framlington Place, Newcastle upon Tyne NE2 4HH, UK.

\*These authors contributed equally to this work.



**Figure 1 | Endothelial-cell FAK deletion sensitizes cancer cells to DNA-damaging therapies *in vivo*.** **a–f**, *Pdgfrb-iCre<sup>ER</sup>;Fak<sup>fl/fl</sup>* and control mice were injected subcutaneously with B16F0 or CMT19T tumour cells (day 0), given tamoxifen (Tam.; from day 7 onwards) to generate ECFAK<sup>KO</sup> and ECFAK<sup>WT</sup> mice, respectively, and subsequently treated or not with DNA-damaging therapy. **a, b**, In untreated mice tumour growth did not differ between genotypes. **c, d**, DNA-damaging therapy significantly inhibited tumour growth in ECFAK<sup>KO</sup> mice when compared with ECFAK<sup>WT</sup> controls. Graphs show mean tumour volumes  $\pm$  standard error of the mean (s.e.m.).  $n = 9$  ECFAK<sup>WT</sup> and 15 ECFAK<sup>KO</sup> mice per test. Horizontal bars represent procedure timelines. Dox., doxorubicin; Irrad., irradiation. **e, f**, Representative images of tumours at experimental endpoints. **g–j**, Immunofluorescence staining analysis for endothelial-cell FAK in PECAM-positive blood vessels in human lymphoma sections. **g**, At diagnosis, a reduced percentage of FAK-positive blood vessels correlates with subsequent achievement of complete remission, but an increased percentage of FAK-positive blood vessels correlates with subsequent disease progression. Bar chart shows the mean percentage of FAK-positive blood vessels  $\pm$  s.e.m.  $n = 16$  biopsy samples taken at diagnosis, 7 of which achieved complete remission and 9 of which subsequently progressed after treatment. Blood vessels were counted from triplicate tissue microarray (TMA) samples. **h**, Endothelial-cell FAK expression was significantly higher in relapsed lymphoma when compared with endothelial-cell FAK expression at diagnosis in matched patient samples. Scatter plot shows mean endothelial-cell FAK fluorescence pixel intensity per sample  $\pm$  s.e.m.  $n = 13$  matched patient biopsies. **i, j**, Representative images of human lymphoma taken at diagnosis and relapse (that is, after treatment including doxorubicin) immunostained for PECAM (red), FAK (green) and 4',6-diamidino-2-phenylindole (DAPI; blue). Arrowheads indicate low FAK expression; arrows indicate high FAK expression. Scale bars, 5 mm (**e, f**); 50  $\mu$ m (**j**). \* $P < 0.05$ , \*\* $P < 0.01$ , Student's *t*-test. NS, not significant.

matched samples taken at diagnosis and at relapse, endothelial-cell FAK expression was elevated significantly at relapse when compared with expression levels at diagnosis (Fig. 1h–j). Overall, our data indicate a

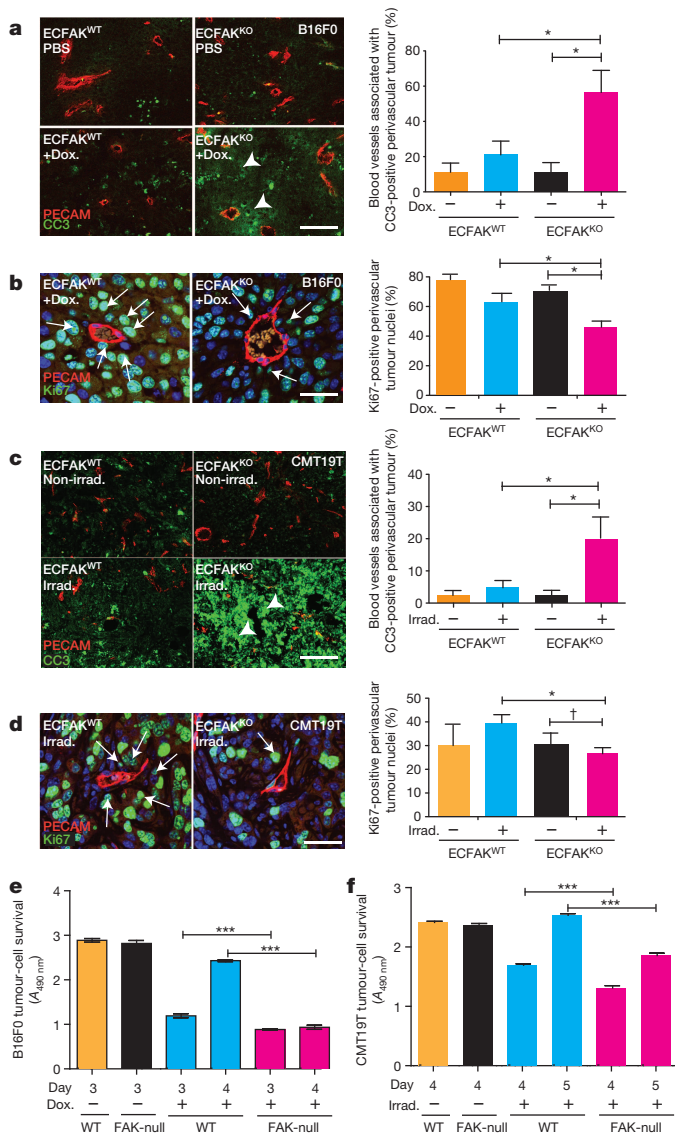
substantial correlation between chemoresistance and endothelial-cell FAK levels in human cancer.

Previous reports have suggested that chemotherapy and/or radiation therapy may result in perivascular chemoresistant niches that actually protect tumour cells from apoptosis<sup>7,19</sup>. However, the regulators of this process within the endothelium *in vivo* have not yet been identified. We show, at 48 h post-treatment cessation, that the number of blood vessels within apoptotic perivascular tumour-cell niches, detected by cleaved caspase 3 staining, was enhanced significantly in doxorubicin-treated ECFAK<sup>KO</sup> mice when compared with similarly treated ECFAK<sup>WT</sup> or placebo-treated mice (Fig. 2a). Furthermore, tumour-cell proliferation, detected by Ki67 staining, was reduced in perivascular zones of doxorubicin-treated ECFAK<sup>KO</sup> mice when compared with controls (Fig. 2b). Similar results were observed for radiotherapy-treated ECFAK<sup>KO</sup> mice (Fig. 2c, d). No differences between ECFAK<sup>WT</sup> and ECFAK<sup>KO</sup> mice in non-treated groups were observed (Fig. 2a–d). These results suggest that, upon DNA damage, endothelial cells may provide protective paracrine signals to tumour cells, which are absent when endothelial-cell FAK is deleted. To confirm this, we show that although conditioned media from untreated wild-type and FAK-null endothelial cells has no apparent effect on tumour-cell survival, conditioned media from either doxorubicin or irradiated wild-type endothelial cells protects cultured tumour cells from DNA damage over time and allows for tumour-cell growth. In contrast, conditioned media from either doxorubicin-treated or irradiated FAK-null endothelial cells confer chemo- and radio-sensitivity to tumour cells, reducing their survival *in vitro* (Fig. 2e, f). Together, these results demonstrate a novel role for endothelial-cell FAK in tumour-cell sensitization to doxorubicin treatment or radiotherapy by the release of paracrine signals.

We next sought to identify the molecular basis for these endothelial-cell effects by FAK. Activation and nuclear translocation of the NF- $\kappa$ B family of transcription factors is known to mediate cellular responses to DNA-damaging therapies<sup>20</sup>; however, a role for FAK-dependent NF- $\kappa$ B functions in endothelial-cell responses to chemotherapy has not been defined previously. In NF- $\kappa$ B luciferase reporter assays, doxorubicin-treated FAK-null endothelial cells exhibited significantly reduced levels of NF- $\kappa$ B activity when compared with similarly treated wild-type controls cells (Fig. 3a). Corroborating these results, while doxorubicin-treated wild-type endothelial cells induced nuclear translocation of the p65 subunit of NF- $\kappa$ B, this was significantly blocked in FAK-null endothelial cells at 4, 24 and 48 h after doxorubicin treatment (Fig. 3b, c and Extended Data Fig. 6). These data are supported by increased levels of phosphorylated p65 (Ser 536) in nuclear fractions from wild-type, but not FAK-null, endothelial cells after doxorubicin treatment (Fig. 3d) and decreased levels of phosphorylated-I $\kappa$ B $\alpha$  in cytosolic fractions of doxorubicin-treated FAK-null endothelial cells (Extended Data Fig. 7). Moreover, nuclear localization of p65 *in vivo*, an indicator of NF- $\kappa$ B activity, was evident *in vivo* in the tumour endothelium of ECFAK<sup>WT</sup>, but not ECFAK<sup>KO</sup>, mice that had been treated with doxorubicin (Fig. 3e). Thus, our data indicate a novel role for endothelial-cell FAK in doxorubicin-induced control of NF- $\kappa$ B activation *in vitro* and *in vivo*.

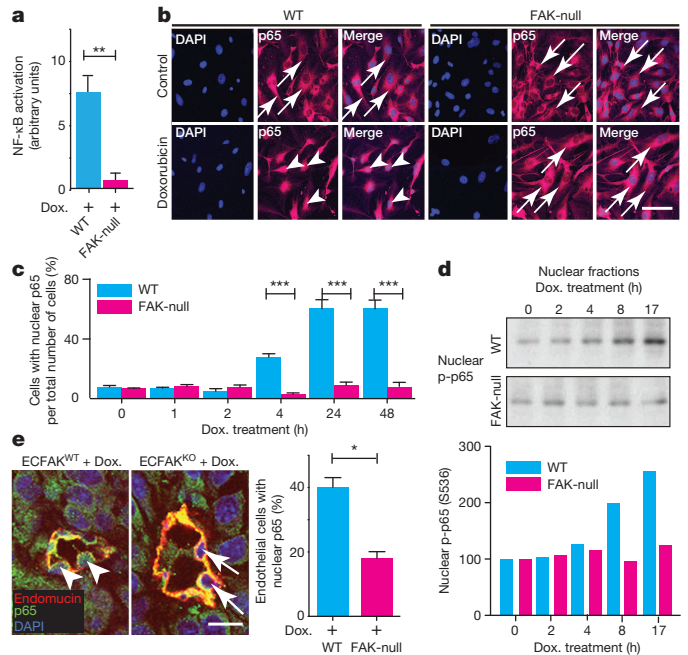
Although other signalling pathways may be involved, FAK-dependent regulation of the NF- $\kappa$ B pathway has been shown to be a major regulator of cytokine production in tumour cells<sup>21–26</sup>. However, the regulation of cytokine production in endothelial cells, especially after DNA-damaging treatment, is not known. Therefore, we next examined whether endothelial-cell FAK deficiency affected doxorubicin-induced cytokine production. Cytokine protein array analysis revealed that doxorubicin stimulation induced an increase in production of several cytokines in wild-type endothelial cells when compared with untreated controls. In contrast, doxorubicin-induced responses were not increased after DNA-damaging therapy in FAK-null endothelial cells (Fig. 4a). A reduction in fold-increase of cytokine levels, some similar to those found after doxorubicin treatment, was also observed after irradiation of FAK-null endothelial cells when compared with irradiated wild-type controls (Extended Data Fig. 8). Cytokine levels were similar between untreated





**Figure 2 | Loss of endothelial-cell FAK sensitizes tumour cells to DNA-damaging therapy *in vivo* and *in vitro*.** **a–d**, Double immunostaining of B16F0 (**a**, **b**) and CMT19T (**c**, **d**) tumour sections from mice treated or not with doxorubicin (Dox.) or irradiation (Irrad.) for the apoptotic marker cleaved caspase 3 (CC3; green; **a**, **c**) or the proliferation marker Ki67 (green; **b**, **d**), and the endothelial marker PECAM (red). DAPI (blue) provides a nuclear marker. Bar charts show quantitation at 48 h post-treatment cessation of the mean number of blood vessels that are within CC3-positive tumour cell niches + s.e.m. (**a**, **c**;  $n = 3$  mice per group) or the percentage of Ki67-positive perivascular tumour cells + s.e.m. (**b**, **d**;  $n = 5$  mice per group). **e**, Conditioned media from untreated (–) and doxorubicin-treated (+) endothelial cells were applied to B16F0 cell cultures and tumour-cell survival was measured.  $n = 9$  technical replicates. **f**, Conditioned medium from non-irradiated (–) or irradiated (+) endothelial cells was applied to irradiated CMT19T cells and tumour-cell survival was measured in MTS assays at 4 and 5 days. Bar charts show mean tumour-cell survival + s.e.m. according to corrected absorbance readings of MTS assays.  $n = 12$  technical replicates.  $A_{490 \text{ nm}}$ , absorbance at 490 nm. WT, wild type. Arrowheads indicate CC3-positive perivascular tumour cells; arrows indicate Ki67-positive perivascular tumour cells. Scale bars, 100  $\mu\text{m}$  (**a**, **c**); 50  $\mu\text{m}$  (**b**, **d**). † $P = 0.1$ , \* $P < 0.05$ , \*\*\* $P < 0.001$ , Student's *t*-test.

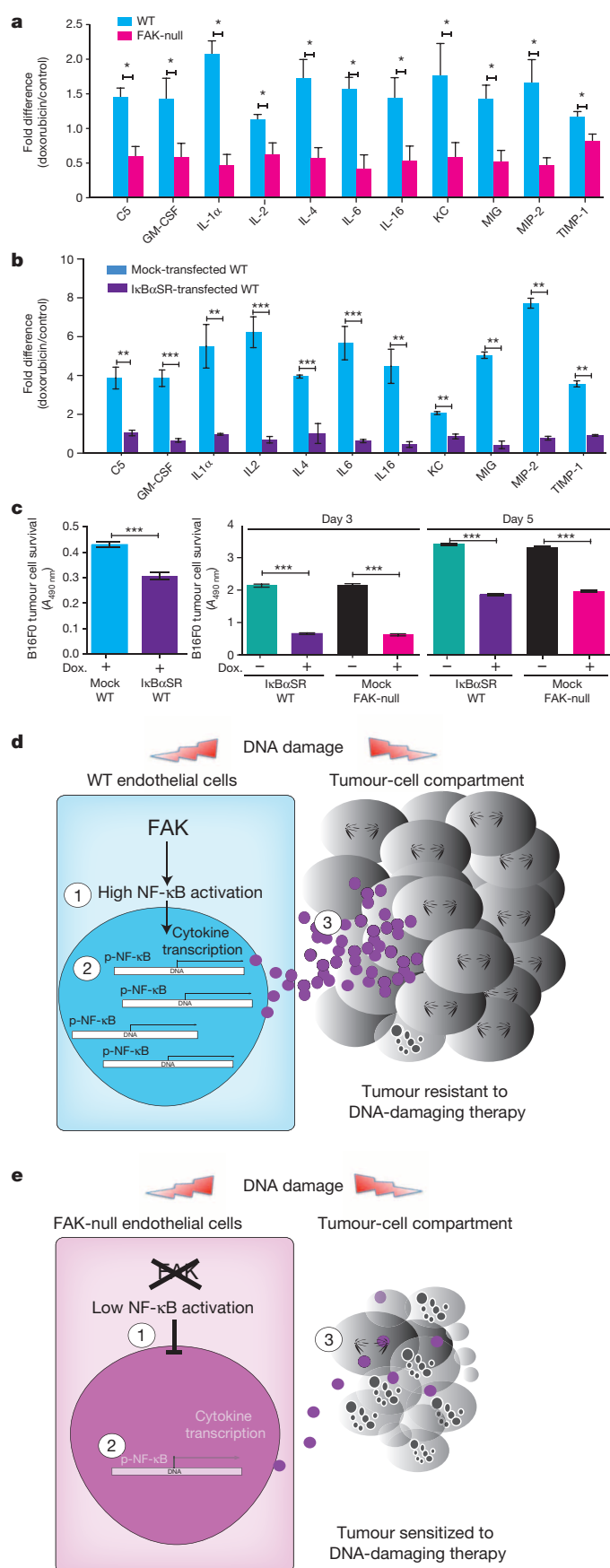
wild-type and FAK-null endothelial cells (Extended Data Fig. 9a). An indicator of cytokine and interleukin activity is phosphorylation of the transcription factor STAT3. We show that doxorubicin treatment enhances the percentage of tumour cells with phosphorylated (p)-STAT3 in ECFAK<sup>WT</sup> mice, but not tumour cells of similarly treated ECFAK<sup>KO</sup>



**Figure 3 | FAK deficiency inhibits doxorubicin-stimulated endothelial-cell p65 activity, phosphorylation and nuclear translocation.** **a**, Luciferase assays indicate that NF-κB activity is significantly reduced in doxorubicin-stimulated FAK-null endothelial cells.  $n = 4$  experimental repeats. Dox., doxorubicin. **b**, Immunofluorescence detection of p65 (red), DAPI (blue) in wild-type (WT) and FAK-null immortalized endothelial cells with and without doxorubicin treatment. Arrows indicate cytoplasmic p65; arrowheads indicate nuclear p65. **c**, Bar chart shows mean percentage of endothelial cells with nuclear p65 ± s.e.m.  $n = 188$ –385 cells per group. **d**, Western blot analysis of nuclear fractions of wild-type and FAK-null endothelial cells after doxorubicin treatment. Bottom, bar chart shows mean densitometry readings over time. p-p65, phosphorylated p65.  $n = 2$  experimental repeats. **e**, Tumour sections from doxorubicin-treated ECFAK<sup>WT</sup> and ECFAK<sup>KO</sup> mice were immunostained for endomucin (red), p65 (green) and DAPI (blue) and the percentage of endothelial cells with nuclear p65 was assessed. Right, bar chart shows mean percentage of endothelial cells with nuclear p65 *in vivo* + s.e.m.  $n = 34$ –89 endothelial cells per tumour and 4 tumours per group. Scale bars, 50  $\mu\text{m}$  (**b**); 20  $\mu\text{m}$  (**e**). \* $P < 0.05$ , \*\* $P < 0.02$ , \*\*\* $P < 0.01$ , Student's *t*-test.

mice (Extended Data Fig. 9b), corroborating our results of reduced cytokine production in doxorubicin-treated ECFAK<sup>KO</sup> mice. These data provide *in vivo* evidence for decreased cytokine effects in doxorubicin-treated ECFAK<sup>KO</sup> when compared with similarly treated ECFAK<sup>WT</sup> mice.

To expand the mechanistic basis of chemosensitization in ECFAK<sup>KO</sup> mice, we tested whether inactivation of the NF-κB signalling pathway in wild-type endothelial cells is sufficient to mimic FAK-null endothelial cells. Wild-type endothelial cells were transfected with the super-repressor, non-phosphorylatable mutant form of IκBα, IκBαSR, which has been shown previously to inhibit the NF-κB pathway<sup>27</sup>. Inhibition of the NF-κB pathway by IκBαSR transfection of wild-type endothelial cells reduced DNA-damage-induced cytokine production when compared with similarly treated mock-transfected controls (Fig. 4b). These data indicate that inhibition of the NF-κB pathway is sufficient to mimic the reduced NF-κB signalling effect in FAK-null endothelial cells after DNA damage. Indeed, conditioned media from doxorubicin-treated, IκBαSR-transfected wild-type endothelial cells was sufficient to sensitize cultured tumour cells to doxorubicin, when compared with mock-transfected wild-type endothelial cells *in vitro*. Furthermore, conditioned medium from doxorubicin-treated IκBαSR-transfected wild-type endothelial cells was able to reduce tumour-cell survival, phenocopying the chemosensitization responses conferred by FAK-null endothelial cells (Fig. 4c). Lastly, intratumoral administration of recombinant



granulocyte-macrophage colony-stimulating factor (GM-CSF) (15 ng) or interleukin (IL)-6 (3 ng), were sufficient to induce similar tumour growth rates in ECFAK<sup>WT</sup> and ECFAK<sup>KO</sup> mice, reversing the chemosensitization

**Figure 4 | Loss of endothelial-cell FAK inhibits doxorubicin-induced, NF-κB-dependent production of endothelial cytokines.** **a**, Quantitation of fold difference in cytokine expression between doxorubicin-treated and non-treated wild-type (WT) and FAK-null endothelial cells + s.e.m.  $n = 3$  experimental repeats. **b**, Quantitation of the fold difference in cytokine expression between doxorubicin-treated and non-treated mock- and IκBαSR-transfected wild-type endothelial cells ± s.e.m.  $n = 4$  experimental repeats. **c**, Conditioned media from doxorubicin-treated mock- or IκBαSR-transfected wild-type endothelial cells were applied to doxorubicin-treated B16F0 cells and cell survival was measured. Conditioned medium from doxorubicin-treated, IκBαSR-transfected wild-type endothelial cells mimics the effects of conditioned medium from doxorubicin-treated mock-transfected FAK-null endothelial cells. Bar charts show mean tumour-cell survival as corrected absorbance readings from MTS, one-step cell survival assays + s.e.m.  $n = 10$  technical repeats. **d**, **e**, Schematic representation of the role of endothelial-cell FAK in tumour-cell sensitization to DNA-damaging therapy. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , Student's  $t$ -test.

phenotype in ECFAK<sup>KO</sup> mice (Extended Data Fig. 10). Together, our results provide proof-of-principle that a decrease in endothelial-cell FAK and a subsequent decrease in DNA-damage-induced NF-κB-dependent endothelial-cell cytokine production controls tumour-cell chemosensitization.

Overall, our data indicate that, upon DNA damage, loss of endothelial-cell FAK is sufficient to sensitize tumour cells to chemotherapy by suppressing NF-κB activation and subsequent cytokine production (Fig. 4d, e). These data establish a new concept in the regulation of chemoresistance. Specifically, our data point to a role for endothelial-cell FAK in the regulation of chemotherapy responses, and provide a starting point for the development of new approaches to improve response to DNA-damaging therapies by specifically targeting endothelial-cell FAK.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 18 April 2013; accepted 29 May 2014.

Published online 27 July 2014.

- De Vita, V. T., Hellman, S. & Rosenberg, S. A. *Cancer: Principles and Practice of Oncology* (Lippincott Williams and Wilkins, 2001).
- Rottenberg, S. & Jonkers, J. Modeling therapy resistance in genetically engineered mouse cancer models. *Drug Resist. Updat.* **11**, 51–60 (2008).
- Straussman, R. *et al.* Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion. *Nature* **487**, 500–504 (2012).
- Wilson, T. R. *et al.* Widespread potential for growth-factor-driven resistance to anticancer kinase inhibitors. *Nature* **487**, 505–509 (2012).
- Sun, Y. *et al.* Treatment-induced damage to the tumour microenvironment promotes prostate cancer therapy resistance through WNT16B. *Nature Med.* **18**, 1359–1368 (2012).
- Nakasone, E. S. *et al.* Imaging tumor-stroma interactions during chemotherapy reveals contributions of the microenvironment to resistance. *Cancer Cell* **21**, 488–503 (2012).
- Gilbert, L. A. & Hemann, M. T. DNA damage-mediated induction of a chemoresistant niche. *Cell* **143**, 355–366 (2010).
- Acharyya, S. *et al.* A CXCL1 paracrine network links cancer chemoresistance and metastasis. *Cell* **150**, 165–178 (2012).
- Lu, J. *et al.* Endothelial cells promote the colorectal cancer stem cell phenotype through a soluble form of Jagged-1. *Cancer Cell* **23**, 171–185 (2013).
- Mitra, S. K. & Schlaepfer, D. D. Integrin-regulated FAK-Src signaling in normal and cancer cells. *Curr. Opin. Cell Biol.* **18**, 516–523 (2006).
- Lim, S. T. *et al.* Nuclear-localized focal adhesion kinase regulates inflammatory VCAM-1 expression. *J. Cell Biol.* **197**, 907–919 (2012).
- McLean, G. W. *et al.* Specific deletion of focal adhesion kinase suppresses tumor formation and blocks malignant progression. *Genes Dev.* **18**, 2998–3003 (2004).
- Shibue, T. & Weinberg, R. A. Integrin β1-focal adhesion kinase signaling directs the proliferation of metastatic cancer cells disseminated in the lungs. *Proc. Natl Acad. Sci. USA* **106**, 10290–10295 (2009).
- Tavora, B. *et al.* Endothelial FAK is required for tumour angiogenesis. *EMBO Mol. Med.* **2**, 516–528 (2010).
- Nakamura, J. *et al.* Biphasic function of focal adhesion kinase in endothelial tube formation induced by fibril-forming collagens. *Biochem. Biophys. Res. Commun.* **374**, 699–703 (2008).
- Stokes, J. B. *et al.* Inhibition of focal adhesion kinase by PF-562,271 inhibits the growth and metastasis of pancreatic cancer concomitant with altering the tumor microenvironment. *Mol. Cancer Ther.* **10**, 2135–2145 (2011).



17. Fisher, R. I. *et al.* Comparison of a standard regimen (CHOP) with three intensive chemotherapy regimens for advanced non-Hodgkin's lymphoma. *N. Engl. J. Med.* **328**, 1002–1006 (1993).
18. Hagemeister, F. B. Treatment of relapsed aggressive lymphomas: regimens with and without high-dose therapy and stem cell rescue. *Cancer Chemother. Pharmacol.* **49** (suppl. 1), 13–20 (2002).
19. Hambardzumyan, D. *et al.* PI3K pathway regulates survival of cancer stem cells residing in the perivascular niche following radiation in medulloblastoma *in vivo*. *Genes Dev.* **22**, 436–448 (2008).
20. Perkins, N. D. The diverse and complex roles of NF- $\kappa$ B subunits in cancer. *Nature Rev. Cancer* **12**, 121–132 (2012).
21. Zhang, H. M. *et al.* Induced focal adhesion kinase expression suppresses apoptosis by activating NF- $\kappa$ B signaling in intestinal epithelial cells. *Am. J. Physiol. Cell Physiol.* **290**, C1310–C1320 (2006).
22. Tseng, W. P., Su, C. M. & Tang, C. H. FAK activation is required for TNF- $\alpha$ -induced IL-6 production in myoblasts. *J. Cell. Physiol.* **223**, 389–396 (2010).
23. Petzold, T. *et al.* Focal adhesion kinase modulates activation of NF- $\kappa$ B by flow in endothelial cells. *Am. J. Physiol. Cell Physiol.* **297**, C814–C822 (2009).
24. Funakoshi-Tago, M. *et al.* Tumor necrosis factor-induced nuclear factor  $\kappa$ B activation is impaired in focal adhesion kinase-deficient fibroblasts. *J. Biol. Chem.* **278**, 29359–29365 (2003).
25. Ben-Neriah, Y. & Karin, M. Inflammation meets cancer, with NF- $\kappa$ B as the matchmaker. *Nature Immunol.* **12**, 715–723 (2011).
26. DiDonato, J. A., Mercurio, F. & Karin, M. NF- $\kappa$ B and the link between inflammation and cancer. *Immunol. Rev.* **246**, 379–400 (2012).
27. Pikarsky, E. *et al.* NF- $\kappa$ B functions as a tumour promoter in inflammation-associated cancer. *Nature* **431**, 461–466 (2004).

**Acknowledgements** We thank A. Papachristodoulou, J. Holdsworth and B. Williams for their help with immunostaining and animal husbandry. Also M. Hemann for his critical appraisal of the manuscript. The work was funded by CR-UK (C9218/A12007), AICR (12-1068), Medical Research Council (G0901609), National Cancer Institute (P01 CA95426:JGG); Leukemia Lymphoma Research (11022); and CR-UK PhD studentship (C1443/A9215).

**Author Contributions** The following authors are listed in the author list in alphabetical order: S.B., F.D., I.F., T.L., D.M.L. and P.-P.W. for their equal and combined contribution to the paper. B.T. and K.M.H.-D. designed the experiments. B.T. performed the experiments. L.E.R. did the GM-CSF rescue experiments *in vivo*, vessel perfusion, doxorubicin delivery, p-STAT3 staining and hypoxia assays. S.B. performed some of the tumour growth and treatment experiments, CD45 analysis, human lymphoma staining and irradiated cytokine responses; F.D. carried out the conditioned media experiments and MTS assays and several histological analyses; I.F. conducted the primary endothelial cell assays; T.L. measured endothelial-cell FAK and blood-vessel FAK levels in human lymphoma; D.M.L. did the transfections and nuclear fractionation experiments; P.-P.W. carried out the transfected cell cytokine arrays. G.E. carried out the histology; A.C. and J.G.G. provided human lymphoma tissue sections and advice; A.L., J.H. and N.P. performed the NF- $\kappa$ B activation assays and A.A. carried out the survival analysis. B.T. and K.M.H.-D. wrote the paper with substantial input from the co-authors.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.M.H.-D. (K.Hodivala-Dilke@qmul.ac.uk).

## METHODS

**Mice.** *Pdgfb-iCre<sup>ER</sup>;Fak<sup>fl/fl</sup>* mice<sup>14</sup> were maintained on a mixed C57BL/6J (ref. 14) background, or on a pure C57 background for Eumyc experiments. Both male and female mice were used, aged 6–24 weeks old.

**Tumour growth, doxorubicin and irradiation treatment.** Mouse melanoma (B16F0, ATCC; mycoplasma free) or mouse lung carcinoma (CMT19T, CR-UK Cell Production; mycoplasma free) cells ( $1 \times 10^6$ ) were injected subcutaneously in the flank of *Pdgfb-iCre<sup>ER</sup>;Fak<sup>fl/fl</sup>* mice and wild-type control mice (*Pdgfb-iCre<sup>ER</sup>*; non-floxed or *Fak<sup>fl/fl</sup>*). Simultaneously, animals were given a soy-free diet (Harlan) to reduce oestrogen levels and increase tamoxifen sensitivity. At days 7 and 8 after tumour inoculation, once tumour growth had begun, all mice were injected intraperitoneally (i.p.) with 150  $\mu$ l of 10 mg ml<sup>-1</sup> of tamoxifen (Sigma, T5648) diluted in 10% ethanol in peanut oil (Sigma) to induce endothelial-cell FAK deletion. From day 8 onwards, all animals were fed with tamoxifen-containing diet (TAM400, Harlan). All animals with B16F0 subcutaneous tumours were injected i.p. with 8 mg kg<sup>-1</sup> of doxorubicin (Accord Healthcare) or PBS as a negative control on days 9, 11 and 13 after tumour-cell inoculation. Alternatively, animals with subcutaneous CMT19T tumours were irradiated with 5 Gy of  $\gamma$ -irradiation on day 10 after tumour injection. For both tumour types calliper measurements were taken over time and animals were killed when tumours reached the maximum size allowed by UK Home Office regulations.

**FAK and PECAM staining in human non-Hodgkin lymphoma samples.** Biopsy samples from non-Hodgkin lymphoma patients, either before (that is, at diagnosis) or after doxorubicin-based chemotherapy (that is, from relapsed patients), were analysed. Formalin-fixed samples were de-waxed; rehydrated; blocked in 10% goat serum; incubated in rabbit anti-FAK antibody (Cell Signaling, 3285) and mouse anti-human CD31 (Leica, CD31-1A10-CE-S) in 0.5% goat serum in PBS overnight at 4 °C; incubated in biotinylated anti-rabbit (DAKO) and Alexa 546 anti-mouse (Invitrogen) diluted 1:100 in 0.5% goat serum in PBS; washed in PBS; and finally processed using streptavidin-HRP/fluorescein kit (TSA Fluorescence Systems). The levels of endothelial-cell FAK were quantitated with Image J software and the mean fluorescence intensity of FAK per pixel was measured.

**Immunostaining.** Sections of human lymphoma or mouse tumours were immunostained for endothelial cells using either anti-PECAM antibody (MEC13.3; BD Biosciences, 553370) or rat anti-endomucin (V.7C7; Santa Cruz, SC-65495) in combination with either rabbit anti-FAK antibody (Cell Signaling, 3285), anti-cleaved caspase 3 (Cell Signaling, 9661), rabbit anti-Ki67 (Abcam, AB15580), or rabbit anti-p65 subunit of NF- $\kappa$ B (D14E12; Cell Signaling, 8242S). See later for details.

**Blood vessel and CC3 immunostaining in mouse tumour sections.** Snap-frozen mouse tumour sections were fixed in acetone for 10 min at -20 °C; rehydrated in PBS for 10 min; blocked in 5% goat serum diluted in PBS for 1 h at room temperature; washed once in PBS; incubated overnight at 4 °C with rat anti-mouse PECAM antibody (MEC13.3; BD Biosciences, 553370) or rabbit anti-cleaved caspase 3 (Cell Signaling, 9661), both diluted 1:100 in 0.5% goat serum in PBS; washed three times in PBS; incubated with anti-rat Alexa 546 (Invitrogen) and anti-rabbit Alexa 488 (Invitrogen) for 1 h at room temperature; washed in PBS; and finally mounted in ProLong Gold with DAPI (Invitrogen). Blood vessels in direct contact with cleaved caspase-3-positive tumour cells were quantified as a percentage of total vessels within five fields of view using a  $\times 40$  objective.

**NF- $\kappa$ B and Ki67 staining in mouse tumours.** Sections from fixed tumours were de-waxed and rehydrated in descending concentrations of ethanol. A 3% hydrogen peroxide diluted in methanol incubation was carried out for 15 min at room temperature between the 100% ethanol immersions. Sections were washed in PBS; microwaved in 10 mM Na citrate buffer (pH 6.0) for 20 min; blocked for 1 h at room temperature in 10% goat serum diluted in PBS; and incubated at 4 °C overnight with the primary antibodies rabbit anti-p65 NF- $\kappa$ B (D14E12; Cell Signaling, 8242S) or rabbit anti-Ki67 (Abcam, AB15580), together with the vessel marker rat anti-endomucin (V.7C7; Santa Cruz, SC-65495), both diluted 1:100 in 1% goat serum in PBS; washed three times in PBS; incubated for 1 h at room temperature with secondary fluorescent antibodies (goat anti-rat Alexa 546 and goat anti-rabbit Alexa 488); washed in PBS and finally mounted in ProLong Gold with DAPI (Invitrogen). Ki67-positive tumour cells, within a perivascular distance of 50  $\mu$ m from PECAM-positive blood vessels, were quantified as a percentage of perivascular DAPI-positive nuclei.

Microscopy was carried out on a LSM510META or LSM710META confocal microscope (Zeiss).

**Endothelial-cell culture.** Primary mouse lung endothelial cells (MLECs) were isolated from *Pdgfb-iCre<sup>ER</sup>;Fak<sup>fl/fl</sup>* or *Pdgfb-iCre<sup>ER</sup>*; non-floxed adult mice as described previously<sup>28</sup>. After a negative sort with rat anti-CD16/CD32 (Serotec, MCA2305EL), cells were immortalized with polyoma middle T (PmT) virus by incubating them over 2 consecutive days for 4 h with supernatant from GpP+E packaging cells<sup>29</sup>. Cells were grown in Mouse Lung Endothelial Cell Media<sup>28</sup> supplemented with 500 nM 4-hydroxytamoxifen (4-OHT). Two positive sorts using rat anti-ICAM2 (Serotec, MCA 2295EL) and sheep anti-rat IgG magnetic beads (Dynabeads) were

carried out as described previously<sup>28</sup>. Tamoxifen-treated endothelial cells isolated from *Pdgfb-iCre<sup>ER</sup>;Fak<sup>fl/fl</sup>* mice gave rise to FAK-depleted endothelial cells (FAK-null) and those isolated from *Pdgfb-iCre<sup>ER</sup>*; non-floxed mice gave rise to FAK wild-type endothelial cells.

**Doxorubicin-induced cytokine production *in vitro*.** Wild-type and FAK-null endothelial cells were treated with 7.5  $\mu$ g ml<sup>-1</sup> mitomycin C (Roche) for 2 h. After plating equal numbers of endothelial cells, cells were cultured for 24 h in full MLEC medium and then the medium was changed for MLEC medium supplemented with 500 nM 4-OHT with, or without, 0.125  $\mu$ M doxorubicin to generate conditioned medium, which was collected after 48 h. The conditioned media from wild-type and FAK-null endothelial cells were harvested and filtered through a 0.22  $\mu$ m filter to remove cell debris before being added to B16F0 tumour cells, plated in 96-well plates. MTS assessment of cell survival was carried out at 3 and 4 days as described later.

**Irradiation-induced cytokine production *in vitro*.** Wild-type and FAK-null endothelial cells were treated with 7.5  $\mu$ g ml<sup>-1</sup> mitomycin C (Roche) for 2 h. Equal numbers of endothelial cells were plated in full MLEC medium for 24 h and irradiated with 5 Gy of X-rays using a RS2000 biological irradiator. Non-irradiated wild-type and FAK-null endothelial cells were used as controls. Cells were cultured for a further 72 h with MLEC medium supplemented with 500 nM 4-OHT to generate conditioned medium. The conditioned media from wild-type and FAK-null endothelial cells were harvested and filtered through a 0.22  $\mu$ m filter to remove cell debris. Three-thousand CMT19T cells per well were plated in 96-well plates and were irradiated with 5 Gy. Endothelial cells conditioned medium was applied to CMT19T tumour cells 8 h after irradiation. Cells underwent MTS assessment of cell survival at days 4 and 5 as described later.

**Conditioned media production from IkB $\alpha$ SR-transfected cells.** Wild-type endothelial cells were transfected using the Nucleofector electroporation system (Lonza) with either the empty vector (mock control) or IkB $\alpha$ SR. The next day, cells were treated with 7.5  $\mu$ g ml<sup>-1</sup> mitomycin C for 2 h (Sigma). Equal numbers of cells were plated in complete medium supplemented with 4-OHT. After 24 h the medium was replaced by fresh complete medium supplemented with 4-OHT  $\pm$  0.125  $\mu$ M doxorubicin. This medium was collected after 48 h incubation at 37 °C and applied to B16F0 cells and tumour-cell survival was assessed using the MTS assay as described later.

**MTS assay for tumour-cell survival.** *In vitro* chemosensitivity was assessed using the CellTiter 96 AQueous One Solution Reagent (Promega). Assays were done by incubating each well, containing tumour cells, with 20  $\mu$ l of reagent in 100  $\mu$ l OptiMEM (Invitrogen) for 90 min. Plates were read at 490 nm, with absorbance corrected relative to blank wells containing reagent only.

**Immunostaining endothelial cells.** Wild-type or FAK-null endothelial cells ( $4 \times 10^4$ ) were grown for 48 h in MLEC medium supplemented with 500 nM 4-OHT with or without 0.125  $\mu$ M doxorubicin. Cells were then serum starved for 4 h in OptiMEM (Invitrogen) with or without 0.125  $\mu$ M of doxorubicin.

For NF- $\kappa$ B detection cells were fixed with 4% paraformaldehyde (PFA) for 20 min at room temperature; washed in PBS three times; and permeabilized in 0.5% NP40 in PBS for 10 min at room temperature; blocked with 0.1% BSA/0.2% Triton X-100 for 10 min at room temperature; washed three times in PBS; incubated for 1 h at room temperature in anti-p65 NF- $\kappa$ B antibody (Cell Signaling); washed three times in PBS; incubated with Alexa-546-conjugated anti-rabbit diluted 1:100 in PBS; washed three times in PBS; and, finally, mounted in Prolong Gold with DAPI.

**Cytokine arrays.** Wild-type and FAK-null endothelial cells or transfected wild-type endothelial cells were grown in normal MLEC media supplemented or not with 0.125  $\mu$ M doxorubicin (Accord Healthcare) or irradiated (5 Gy). After 48 h of serum starvation, whole-cell lysates were extracted at 48 h (for doxorubicin-treated endothelial cells) and 72 h (for irradiated endothelial cells). Mouse cytokine arrays (Proteome Profiler ARY006, R&D Systems) were processed according to the manufacturer's instructions using 100  $\mu$ g of lysates (in 3% SDS, 60 mM sucrose, 65 mM Tris-HCl pH 6.8) per membrane. Pixel analysis was used for quantification with Image J software.

**Western blot analysis of nuclear fraction p65.** Wild-type and FAK-null endothelial cells were serum starved for 18 h in OptiMEM supplemented with 2% FCS and tamoxifen. The cells were stimulated with 0.125  $\mu$ M doxorubicin in the same medium without tamoxifen. Cells were scraped in hypotonic buffer and left on ice for 5 min. After centrifugation for 10 min at 500g the supernatants were collected as cytosolic extracts. The pellets were washed once in hypotonic buffer, then resuspended and sonicated in nuclear extraction buffer and centrifuged to remove debris. The following antibodies were used for western blot analysis: anti-phospho-p65 NF- $\kappa$ B (S536; Cell Signaling, 3033), anti-p65 NF- $\kappa$ B (D14E12; Cell Signaling, 8242).

**NF- $\kappa$ B activation assay.** Wild-type and FAK-null cell lines were co-transfected with 2  $\mu$ g each of the 3  $\times$  kb ConA NF- $\kappa$ B reporter firefly luciferase plasmid and an internal control plasmid expressing *Renilla* luciferase (pRL-TK, Promega), using the Basic Endothelial Cell kit and T23 program on the Nucleofector II (Lonza) according to the manufacturer's instructions. After 24 h, cells were treated with 0.125  $\mu$ M of doxorubicin, and a further 48 h later cells were actively lysed with Passive Lysis Buffer (Promega) and assayed using the Dual Luciferase Reporter Assay system

(Promega) in a Lumat LB9507 luminometer (Berthold Technologies)<sup>30</sup>. Results shown are the mean plus s.e.m. of at least three independent experiments and were normalized to the expression of the internal control plasmid. Statistical analysis was performed using Prism 5 (GraphPad).

**FAK immunostaining in mouse tumour sections.** Tumours were fixed, paraffin embedded and processed as described earlier. For detection of FAK, a rabbit anti-FAK antibody (Cell Signaling, 3285) was used. For FAK staining, amplification of signal was achieved by incubating samples after the primary antibody incubation with anti-rabbit biotinylated (DAKO, E0353) and anti-rat Alexa 546 (Invitrogen) diluted 1:100 in 1% goat serum in PBS. After three washes in PBS, the sections were stained with a streptavidin-HRP/fluorescein kit (TSA Fluorescence Systems).

**Blood vessel density.** Snap-frozen sections of tumours were immunostained for PECAM to detect blood vessels as described earlier. Blood vessel density was calculated by counting the total number of blood vessels across entire midline sections from age-matched, size-matched tumours. Blood vessel density is presented as the number of blood vessels per mm<sup>2</sup> of tumour section.

**Blood vessel perfusion and permeability.** *Pdgfb-iCre<sup>ER</sup>*;non-floxed mice were injected subcutaneously with  $1 \times 10^6$  B16F0 tumour cells. At 7 and 8 days after tumour inoculation, all mice were injected i.p. with 150  $\mu$ l of 10 mg ml<sup>-1</sup> tamoxifen. From day 8 onwards, all mice were fed with tamoxifen containing diet. On days 9, 11 and 13 after tumour-cell inoculation, mice were injected with PBS. On day 14, mice were injected with 100  $\mu$ l of PE-conjugated anti-PECAM antibody (Biolegend) via the tail vein 10 min before killing the mice to analyse blood vessel perfusion. One minute before mice were killed, they were also injected via the tail vein with Hoechst dye (4  $\mu$ g ml<sup>-1</sup>) to analyse blood vessel permeability. Tumours were dissected and immediately snap-frozen. Cryosections were air-dried then fixed in -20 °C acetone for 10 min. Sections were then rehydrated in PBS and washed once with water and mounted with Prolong Gold. The level of blood vessel perfusion was calculated as the percentage of tumour blood vessels that were positive for PE-PECAM over total blood vessel counts. Permeability was analysed by counting the numbers of Hoechst positive nuclei per field of view.

**$\gamma$ H2AX immunostaining on mouse tumour sections.** For mice used and treatment schedules see earlier. On day 14, tumours were removed and snap-frozen. Cryosections were fixed in acetone for 10 min at -20 °C and rehydrated in PBS for 10 min. Staining was performed using the M.O.M. Fluorescein kit (Vector Labs). Sections were incubated for 1 h at room temperature with anti-mouse phosphohistone  $\gamma$ H2AX (Ser 139) antibody (Millipore, clone JBW301). Nuclei of the tumour cells with  $\gamma$ H2AX foci were quantified as a percentage of total nuclei.

**Doxorubicin delivery.** *Pdgfb-iCre<sup>ER</sup>*; *Fak<sup>fl/fl</sup>* mice and wild-type control mice (*Pdgfb-iCre<sup>ER</sup>*;non-floxed) were given a subcutaneous injection of  $1 \times 10^6$  B16F0 cells, then treated with tamoxifen (as described earlier) to generate ECFAK<sup>KO</sup> and ECFAK<sup>WT</sup> mice. To assess delivery of doxorubicin to the tumours, mice were injected with 20 mg kg<sup>-1</sup> doxorubicin over 1 min, 5 min before euthanasia. Under terminal anaesthesia mice were perfused with 4% paraformaldehyde. Perfused tumours were removed and fixed overnight in 4% paraformaldehyde, then transferred to 70% ethanol. Tumours were embedded in paraffin, sectioned, rehydrated and counterstained with DAPI. Tumour sections were analysed using the Zeiss Axioplan microscope and images were captured using Axiovision Rel.4 software.

**Tumour hypoxia.** For mice used and treatment schedules see earlier. On day 14, 1 h before mice were killed, tumour-bearing mice were injected with 60 mg kg<sup>-1</sup> pimonidazole hydrochloride (HypoxprobeTM-1 HPI, diluted in ddH<sub>2</sub>O to a final concentration of 10 mg ml<sup>-1</sup>) intravenously via the tail vein. Tumours were processed immediately after cervical dislocation. Cryosections were thawed, rehydrated and fixed for 10 min in -20 °C acetone then incubated with 1:10 anti-pimonidazole antibody to identify hypoxic areas. Sections were then washed and mounted with ProLong Gold with Antifade plus DAPI (Invitrogen, P36930). Images were taken with a Zeiss microscope and AxioPlan camera. The total tumour area and pimonidazole-positive areas were measured using Axiovision Rel. 4 software (Zeiss). The percentage hypoxic area for each tumour section was then calculated.

**CD45 infiltration.** For mice used and treatment schedules see earlier. On day 14, tumours were harvested and snap-frozen. Cryosections were fixed in acetone for 10 min at -20 °C; rehydrated in PBS for 10 min; blocked in 5% goat serum diluted in PBS for 1 h at room temperature; washed once in PBS; incubated overnight at 4 °C with rat anti-mouse CD45 antibody (Serotec, MCA1388) diluted 1:100 in 0.5% goat serum in PBS; washed three times in PBS; incubated with anti-rat Alexa 488 (Invitrogen) for 1 h at room temperature; washed in PBS and finally mounted in ProLong Gold with DAPI (Invitrogen). CD45 was quantified as a percentage of total DAPI area within five fields of view using a  $\times 40$  objective.

**Experimental metastasis survival experiments.** For experimental metastasis assays,  $0.5 \times 10^6$  B16F10 and EumycBCL2 cells (obtained from S. Hallam and T. Hagemann) were injected via the tail vein of *Pdgfb-iCre<sup>ER</sup>*; *Fak<sup>fl/fl</sup>* mice and control mice (*Pdgfb-iCre<sup>ER</sup>*;non-floxed). When tumours had grown, at either days 7 and 8 (for mice with B16F10) or days 10 and 11 (for EumycBCL2) after tumour-cell inoculation,

mice were given tamoxifen to induce endothelial-cell FAK deletion in *Pdgfb-iCre<sup>ER</sup>*; *Fak<sup>fl/fl</sup>* but not *Pdgfb-iCre<sup>ER</sup>*;non-floxed mice, generating ECFAK<sup>KO</sup> and ECFAK<sup>WT</sup> mice, respectively. Mice were then treated with placebo or doxorubicin at days 12, 13 and 14 (for mice with B16F10) or days 11, 13 and 15 (for EumycBCL2) after tumour-cell inoculation. Survival was recorded for each animal.

**Phospho-STAT3 immunostaining.** *Pdgfb-iCre<sup>ER</sup>*; *Fak<sup>fl/fl</sup>* mice and wild-type control mice (*Pdgfb-iCre<sup>ER</sup>*;non-floxed) were given a subcutaneous injection of  $1 \times 10^6$  CMT19T tumour cells and the treatment schedule was performed as described earlier. On day 10 after tumour-cell inoculation, mice were irradiated, or not, with 5 Gy  $\gamma$ -irradiation. Tumours were harvested when they had reached the maximum legal size allowed by the UK Home Office regulations, and snap-frozen immediately. Cryosections were permeabilized with ice-cold methanol for 10 min, followed by one PBS wash. Slides were blocked (5% normal goat serum, 0.3% Triton X-100 in PBS) for 60 min at room temperature then incubated with phospho-STAT3 (Tyr705) (Cell Signaling; 1:30 dilution in 1% BSA, 0.3% Triton-X 100 in PBS) primary antibody overnight at 4 °C. The slides were washed three times with PBS and incubated with fluorescent-conjugated secondary antibody (1:100 dilution in 1% BSA, 0.3% Triton-X 100 in PBS). Slides were rinsed in PBS three times and once in water containing DAPI (1:10,000) and then mounted with coverslips using Prolong Gold Anti-fade Reagent.

**Endothelial apoptosis.** Apoptosis was measured in tumour sections by double immunostaining for either cleaved caspase 3 (CC3) or TdT-mediated dUTP nick end labelling (TUNEL) and PECAM. Immunostaining was performed as described earlier. Endothelial apoptosis was calculated by counting the percentage of tumour blood vessels that were also CC3-positive or TUNEL-positive.

**Primary endothelial-cell culture and nuclear translocation of p65.** Wild-type and FAK-null primary endothelial cells were generated as described previously<sup>28</sup>. Wild-type and FAK-null primary endothelial cells were grown for 24 h or 48 h with MLEC supplemented with 500 nM 4-hydroxytamoxifen with or without doxorubicin (0.125  $\mu$ M or 0.25  $\mu$ M). Immunostaining for p65 was performed as described earlier.

**Cultured endothelial-cell FAK expression analysis.** Endothelial cells were prepared as described earlier. For FAK detection, cells were fixed with ice-cold acetone for 5 min at -20 °C. The fixed cells were then blocked for 30 min at room temperature in 5% goat serum in PBS and washed once with PBS, then incubated with mouse anti-FAK antibody (77/FAK; BD Biosciences, 610088) diluted 1:100 in PBS. This was followed by three washes in PBS and the cells were then incubated for 1 h at room temperature with Alexa-488-conjugated anti-mouse (Invitrogen) diluted 1:100 in PBS. After three final washes in PBS the coverslips were mounted in Prolong Gold (Invitrogen).

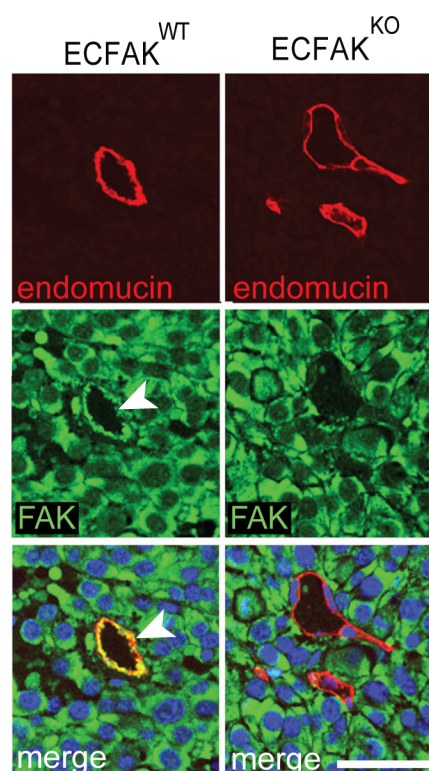
**Western blotting.** Wild-type and FAK-null endothelial cells were grown and lysed under the same conditions as described earlier. Anti-FAK (BD Biosciences) was used to detect FAK expression levels. Western blot analysis was processed with 50  $\mu$ g of total cell lysate as described previously<sup>14</sup>. The following antibodies were used for western blot analysis of cytosolic fractions: phospho-I $\kappa$ B $\alpha$  (S32, Cell Signaling, 2859), I $\kappa$ B $\alpha$  (Santa-Cruz Biotechnology, sc-371).

**GM-CSF and IL-6 intratumoral injections.** B16F0 subcutaneous tumours were grown in *Pdgfb-iCreER*; *Fak<sup>fl/fl</sup>* mice and wild-type control mice (*Pdgfb-iCreER*;non-floxed), and given tamoxifen after tumour growth had begun and treated with doxorubicin as described earlier. On days 9, 10, 11, 12 and 13 after tumour inoculation, half the animals were injected intratumorally with 100  $\mu$ l of either 15 ng ml<sup>-1</sup> or 30 ng ml<sup>-1</sup> of recombinant mouse GM-CSF or IL-6 diluted in PBS (PeproTech)—concentrations that mimic the levels of wild-type endothelial-cell GM-CSF and IL-6 production. Controls were injected with 100  $\mu$ l of PBS. Tumours were measured twice a week and the animals were killed when tumours reached maximum legal size.

**Statistical analysis.** Results are presented as means  $\pm$  s.e.m. for at least 2–3 independent experiments, unless otherwise stated. The sample sizes used were based on level of changes and consistency expected. Statistical significance was reported as appropriate. For animal experiments, animals were excluded from the analysis if tumour volume breached the Home Office legal size limit. Sample sizes were chosen on the basis of the level of changes expected. No randomization methods were used. During animal experiments the investigator was blinded to the genotype of the animals under study. *P* values were calculated with the two-tailed Student's *t*-test unless otherwise stated. *P* < 0.05 was considered statistically significant.

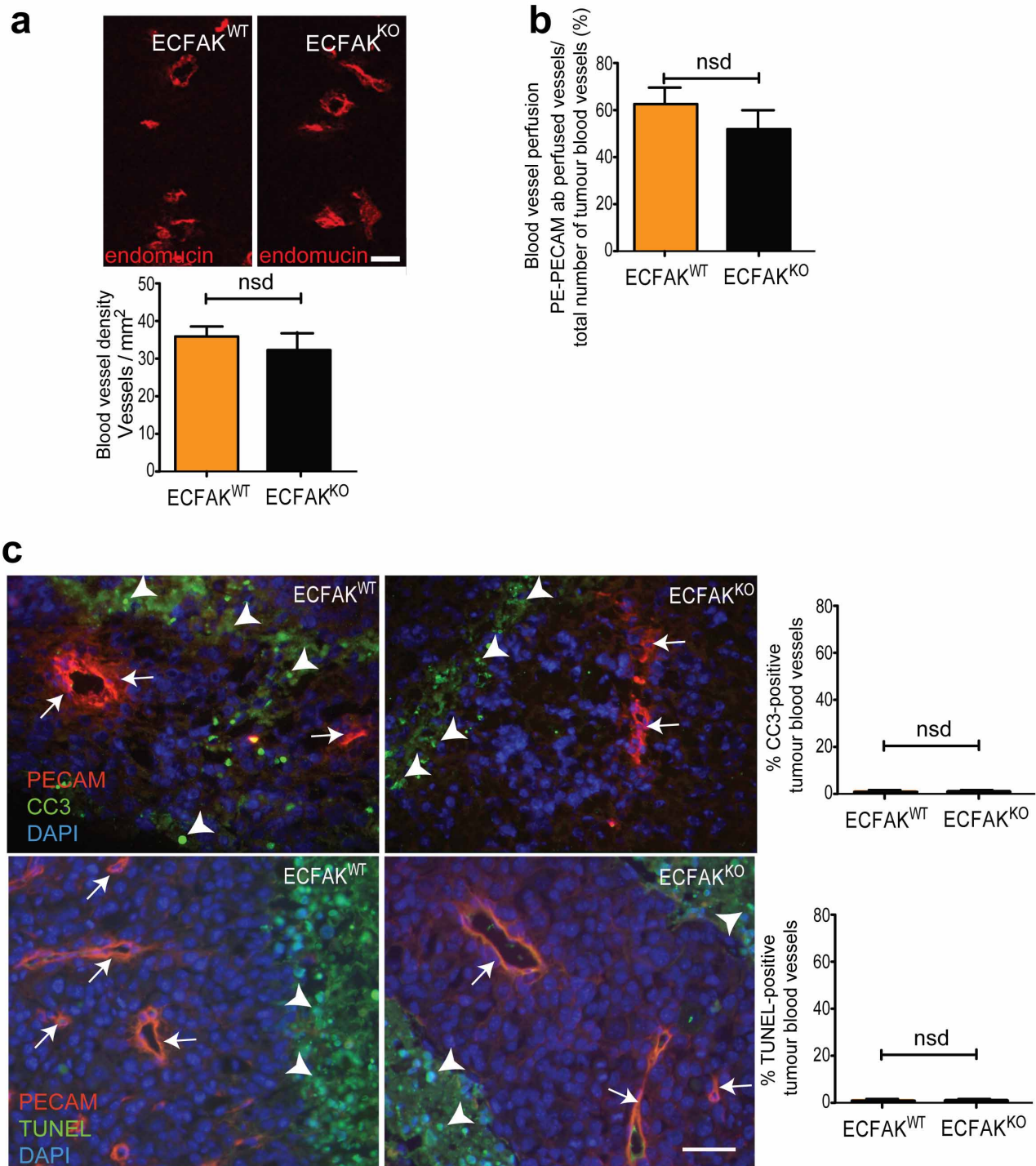
**Ethical regulations.** All animals were used according to the UK Home Office regulations. Human lymphoma samples were obtained with signed informed consent from patients and Ethical Committee approval.

28. Reynolds, L. E. & Hodivala-Dilke, K. M. Primary mouse endothelial cell culture for assays of angiogenesis. *Methods Mol. Med.* **120**, 503–509 (2006).
29. May, T. et al. Establishment of murine cell lines by constitutive and conditional immortalization. *J. Biotechnol.* **120**, 99–110 (2005).
30. Rocha, S., Campbell, K. J. & Perkins, N. D. p53- and Mdm2-independent repression of NF- $\kappa$ B transactivation by the ARF tumor suppressor. *Mol. Cell* **12**, 15–25 (2003).



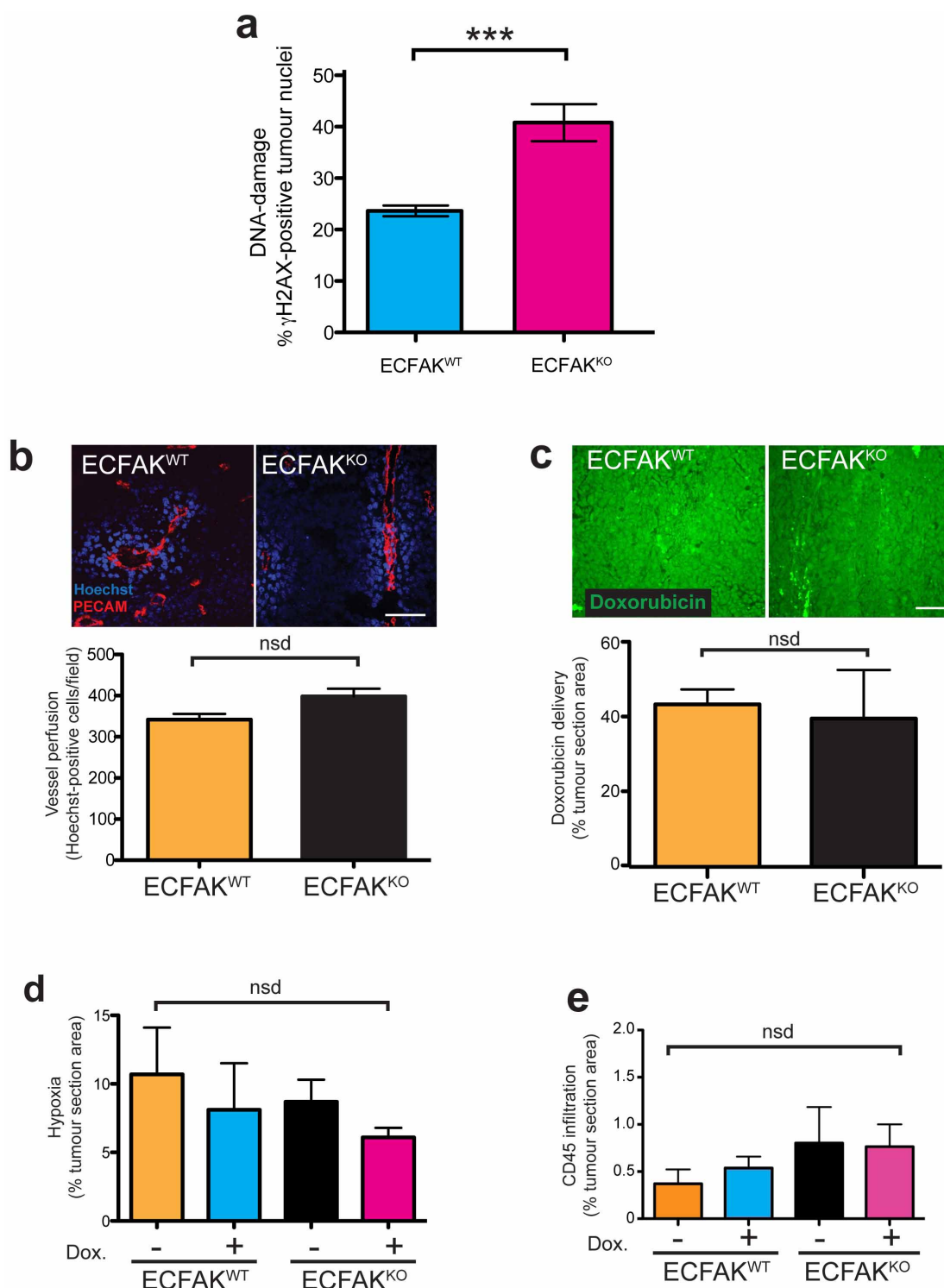
**Extended Data Figure 1 | Loss of endothelial-cell FAK in established tumours.** *Pdgfb-iCreER;Fak<sup>fl/fl</sup>* mice<sup>14</sup> and wild-type control mice (*Pdgfb-iCreER*;non-floxed) were injected subcutaneously with B16F0-melanoma cells. At day 7 after tumour-cell injection, once tumour growth was established, mice were given tamoxifen to induce, or not, endothelial-cell FAK deletion (generating ECFAK<sup>KO</sup> and ECFAK<sup>WT</sup> mice) and tumours continued to grow until they reached the legal size limit at day 24 after tumour-cell injection. Immunofluorescence staining of tumour sections from ECFAK<sup>WT</sup> and ECFAK<sup>KO</sup> mice for FAK (green) and endomucin (red) shows efficient deletion of endothelial FAK in tumour blood vessels when FAK deletion is induced after tumour growth has begun. DAPI staining is shown in blue. Endothelial-specific FAK deletion *in vivo* was confirmed for all experiments. Representative images are given for a minimum of 5 mice per genotype. Scale bar, 75  $\mu$ m.





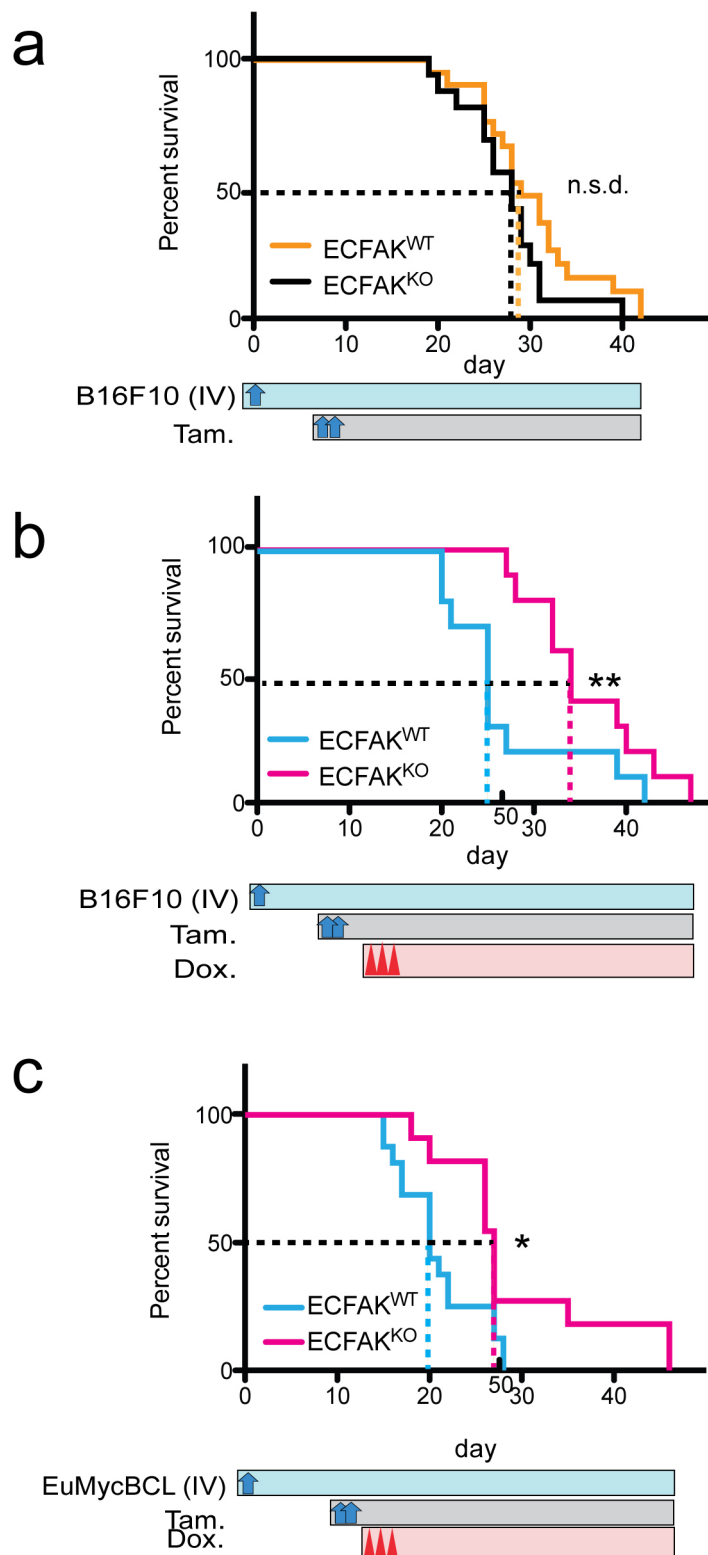
**Extended Data Figure 2 | Loss of endothelial-cell FAK in established tumours does not affect tumour blood vessel density, perfusion or endothelial apoptosis.** **a–c**, *Pdgfb-iCreER;Fak<sup>fl/fl</sup>* mice and wild type control mice (*Pdgfb-iCreER*;non-floxed) were injected subcutaneously with B16F0-melanoma cells. At day 7 after tumour-cell injection, once tumour growth was established, mice were given tamoxifen to induce, or not, endothelial-cell FAK deletion (generating ECFAK<sup>KO</sup> and ECFAK<sup>WT</sup> mice, respectively) and tumours continued to grow until they reached the legal size limit at day 24 after tumour-cell injection. Blood vessels were analysed histologically in midline tumour sections. **a**, Tumour blood vessel density was not affected by the deletion of FAK after tumour growth had begun. Immunofluorescence of endomucin-stained blood vessels and quantitation of number of blood vessels per mm<sup>2</sup> of tumour section are given. **b**, In an ante-mortem procedure, tumour burdened ECFAK<sup>KO</sup> and ECFAK<sup>WT</sup> mice were injected intravenously with

PE-conjugated PECAM antibody. Midline sections were immunostained to detect endomucin-positive vessels. Examination of the percentage of endomucin-positive blood vessels that are PE-PECAM-positive gives an indication of blood vessel perfusion. Tumour blood vessel perfusion was not affected significantly by the deletion of FAK after tumour growth had begun. **c**, Double immunostaining for tumour endothelial cells (PECAM), and either cleaved caspase 3 (CC3) or TUNEL, and DAPI in tumour sections from ECFAK<sup>WT</sup> and ECFAK<sup>KO</sup> mice. Apoptotic tumour cells are clearly visible (arrowheads). In contrast, apoptosis is not detectable in the endothelium of either genotype (arrows). Quantitation of the percentage of CC3-positive or TUNEL-positive tumour endothelial cells showed no significant difference between genotypes. Bar charts show means  $\pm$  s.e.m.  $n = 5–7$  mice per group. NS, not significant, Student's *t*-test. Scale bars, 100  $\mu$ m.



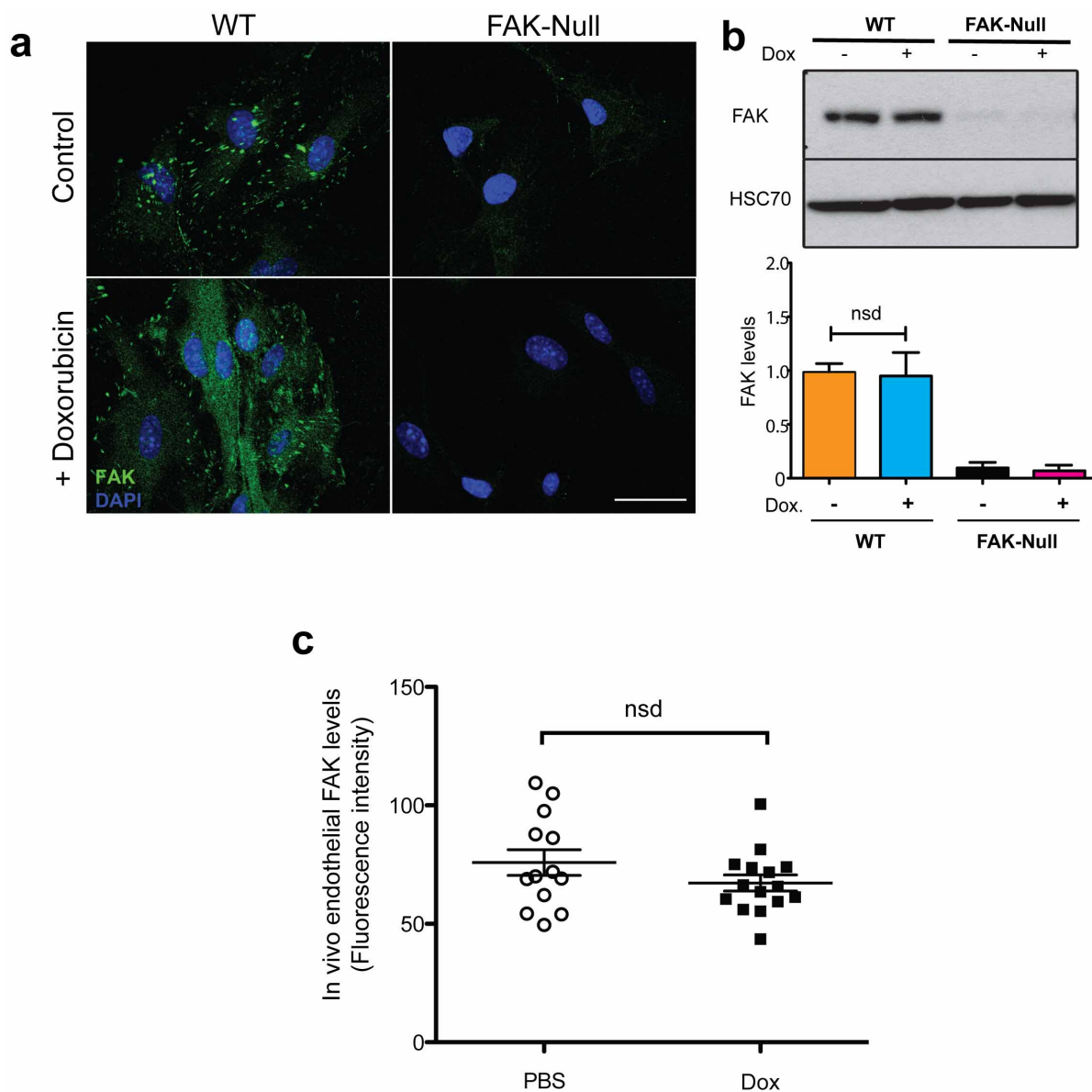
**Extended Data Figure 3 | Increased tumour-cell DNA damage without changes in blood vessel permeability, doxorubicin delivery, hypoxia or CD45 infiltration in ECFAK<sup>KO</sup> mice.** **a**, Quantitation of  $\gamma$ H2AX immunostaining indicates that the level of DNA damage in the tumour-cell compartment is increased in treated ECFAK<sup>KO</sup> mice when compared with ECFAK<sup>WT</sup> mice. Bar chart shows mean percentage of  $\gamma$ H2AX-positive tumour cell nuclei  $\pm$  s.e.m.  $n = 3$  mice per group. **b**, Mice were injected via the tail vein with Hoechst dye and PE-PECAM, in an ante-mortem process, and tumour sections were analysed for blood vessel permeability. Representative images of tumour sections showing Hoechst uptake in tumour cells and PE-PECAM-positive blood vessels are given. Bar chart shows mean number of Hoechst-positive nuclei per field of view for tumours grown in ECFAK<sup>WT</sup> and ECFAK<sup>KO</sup> mice  $\pm$  s.e.m.  $n = 8$  mice per group. **c**, Mice were injected via the tail vein with

doxorubicin (20 mg kg<sup>-1</sup>) and analysed for levels of autofluorescent doxorubicin delivery. Representative images of autofluorescent doxorubicin are given. Bar chart shows mean percentage of doxorubicin-positive area proportional to tumour section area  $\pm$  s.e.m.  $n = 7$  mice per group. **d**, Mice were treated or not with doxorubicin at 9, 11 and 13 days, injected via the tail vein with pimonidazole at day 14 and killed 1 h thereafter. Tumour sections were immunostained to detect hypoxia using an anti-pimonidazole antibody. Bar chart shows percentage hypoxic tumour section area  $\pm$  s.e.m.  $n = 7$  mice per genotype. **e**, Mice were treated or not with doxorubicin as in **c** and tumour sections were immunostained for CD45-positive immune cells. Bar chart shows mean CD45 infiltration as a percentage of CD45-positive cell area over tumour section area  $\pm$  s.e.m.  $n = 4$ –6 mice per group. NS, not significant, Student's *t*-test. Scale bar, 100  $\mu$ m.



**Extended Data Figure 4 | Endothelial-cell FAK deletion enhances the mean survival of doxorubicin-treated mice in experimental metastasis models of melanoma and lymphoma.** **a–c**, *Pdgfb-iCre<sup>ER</sup>;Fak<sup>fl/fl</sup>* mice and wild-type control mice (*Pdgfb-iCre<sup>ER</sup>*;non-floxed) were injected via the tail vein (intravenously) with either B16F10 melanoma cells (**a**, **b**) or EumycBCL2 lymphoma cells (**c**) to establish experimental metastasis models. After tumour growth was established endothelial-cell FAK deletion was induced, or not, by treatment of *Pdgfb-iCre<sup>ER</sup>;Fak<sup>fl/fl</sup>* and *Pdgfb-iCre<sup>ER</sup>*;non-floxed mice with tamoxifen (generating ECFAK<sup>KO</sup> and ECFAK<sup>WT</sup> mice, respectively). Mice were then either treated with placebo (**a**) or doxorubicin (**b**, **c**) and survival of

the mice was recorded. Data show that endothelial-cell FAK deletion, after tumour growth was established, had no effect on survival per se (**a**). In contrast, in both the B16F10 and EumycBCL2 experimental metastasis assays, the deletion of endothelial-cell FAK was sufficient to significantly extend median survival after treatment with doxorubicin (**b**, **c**). Dashed lines represent median survival. Timelines for tamoxifen and doxorubicin treatment are given in the horizontal bars below each graph.  $n = 10$ – $20$  mice per genotype per test. \* $P = 0.0209$ , \*\* $P = 0.0055$ , Gehan–Breslow–Wilcoxon Test. NS, not significant.

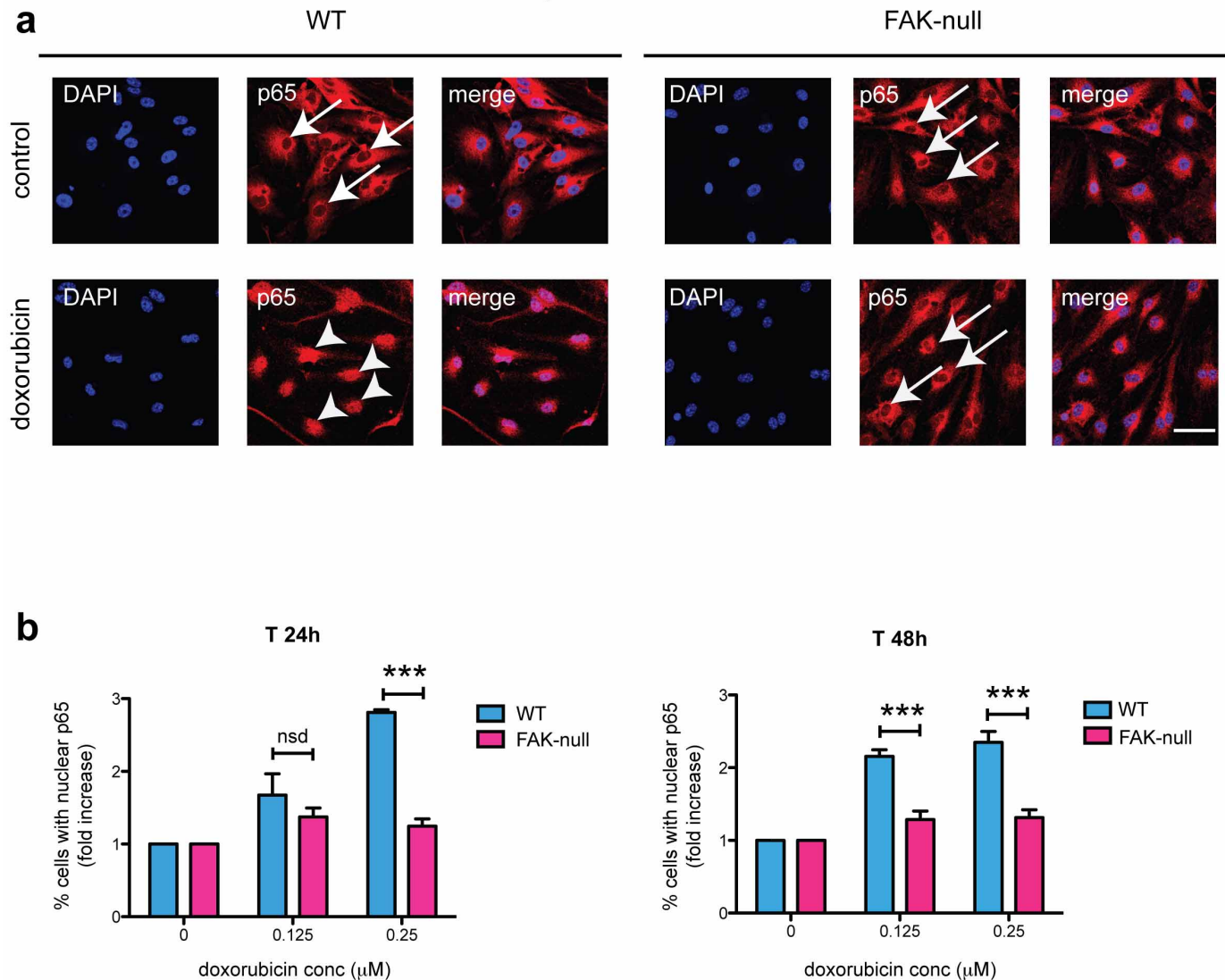


**Extended Data Figure 5 | Alterations in FAK distribution but not expression levels in doxorubicin-stimulated wild-type endothelial cells *in vitro* and *in vivo*.** **a**, Double immunofluorescence staining of cultured wild-type (WT) and FAK-null endothelial cells, with or without doxorubicin treatment, for FAK (green) and DAPI (blue). Staining confirms that FAK is not detectable in FAK-null endothelial cells. In contrast FAK is redistributed in doxorubicin-treated wild-type endothelial cells when compared with untreated wild-type controls. Experiments are representative of three repeats. **b**, Western blot analysis confirms that FAK levels are not significantly changed in

doxorubicin-treated wild-type endothelial cells. HSC70 acts as a loading control. Bar chart represents mean densitometric readings of FAK levels relative to controls.  $n = 3$ . **c**, FAK expression levels are not different after doxorubicin treatment *in vivo*. Image J analysis of endothelial-cell FAK intensity was performed on tumour sections from PBS or doxorubicin-treated ECFAK<sup>WT</sup> mice stained for endomucin and FAK. Graph shows average FAK fluorescence pixel intensity levels in endomucin-positive endothelium for individual mice, with means per group  $\pm$  s.e.m.  $n = 13-15$  mice per treatment group. NS, not significant, Student's *t*-test. Scale bar, 50  $\mu$ m.

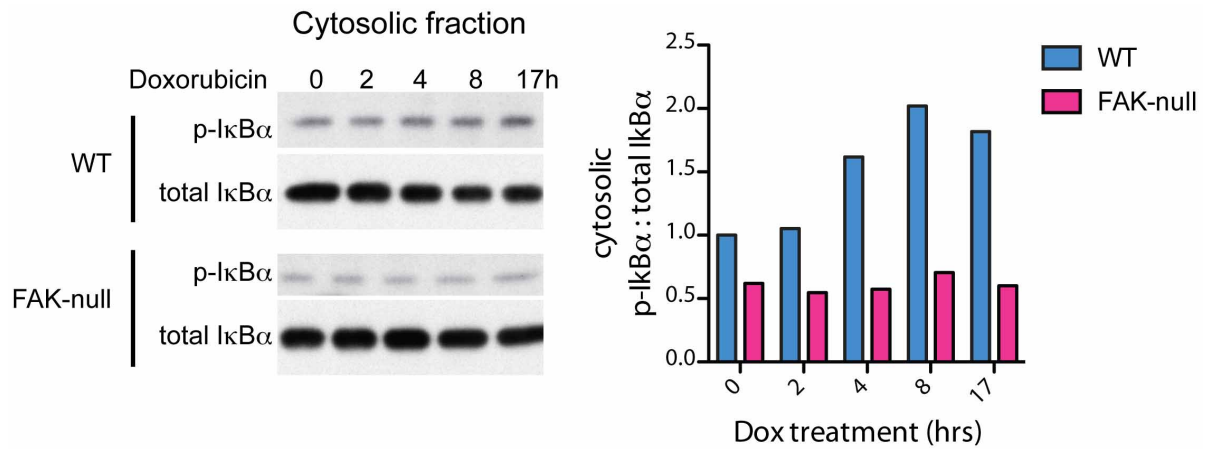


## Primary endothelial cells



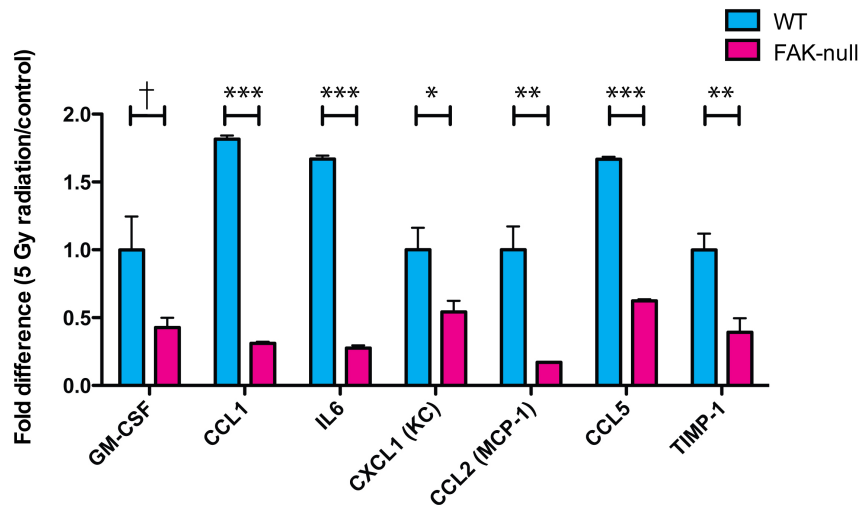
**Extended Data Figure 6 | FAK deficiency inhibits doxorubicin-induced p65 nuclear localization in primary endothelial cells.** **a**, Doxorubicin-induced p65 nuclear translocation is inhibited in FAK-null primary endothelial cells *in vitro*. Wild-type (WT) and FAK-null primary lung endothelial cells were treated for 24 h with tamoxifen and 0.25  $\mu\text{M}$  doxorubicin. Immunostaining for p65 (red) was performed. DAPI staining is shown in blue. Arrows indicate

cytoplasmic p65; arrowheads indicate nuclear p65. Scale bar, 50  $\mu\text{m}$ . **b**, Bar charts show the percentage of endothelial cells with nuclear p65 (fold increase) after doxorubicin treatment (0.125  $\mu\text{M}$  or 0.25  $\mu\text{M}$ ) for 24 h (T24) or 48 h (T48).  $n = 91$ –205 cells per test group. \*\*\* $P < 0.0001$ , Student's  $t$ -test. NS, not significant.



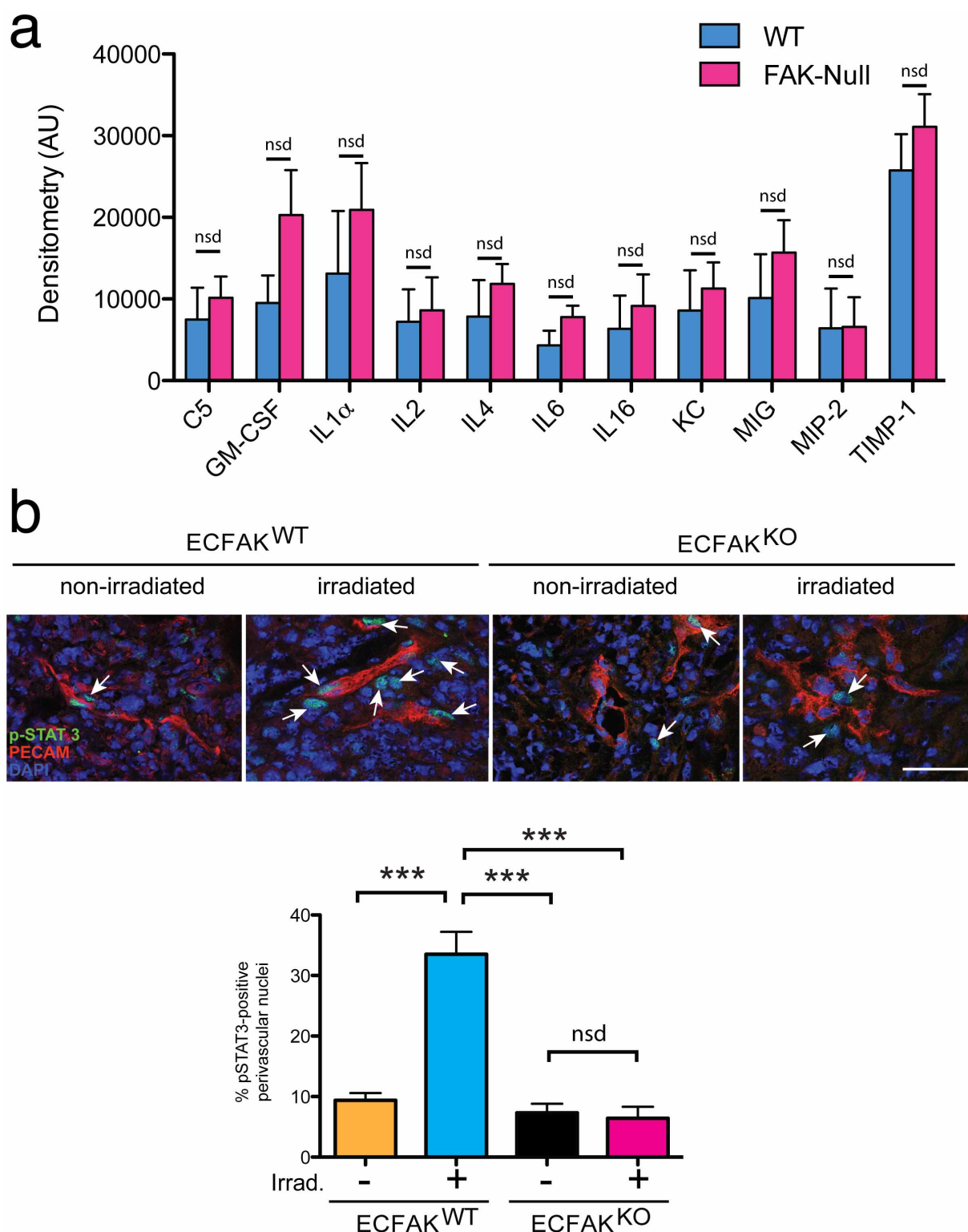
**Extended Data Figure 7 | Phosphorylation of IκBα is reduced in doxorubicin-treated FAK-null endothelial cells.** Wild-type (WT) and FAK-null endothelial cells were treated with 0.125 μM doxorubicin for the time

indicated. Representative western blot of cytosolic phospho-S32-IκBα and total IκBα. Bar chart shows mean densitometric readings from two biological replicates.



**Extended Data Figure 8 | Loss of endothelial-cell FAK inhibits the production of irradiation-induced endothelial cytokines.** Wild-type (WT) and FAK-null endothelial cells were treated with 5 Gy irradiation. Seventy-two hours later, whole-cell lysates were extracted and used in proteome profiler

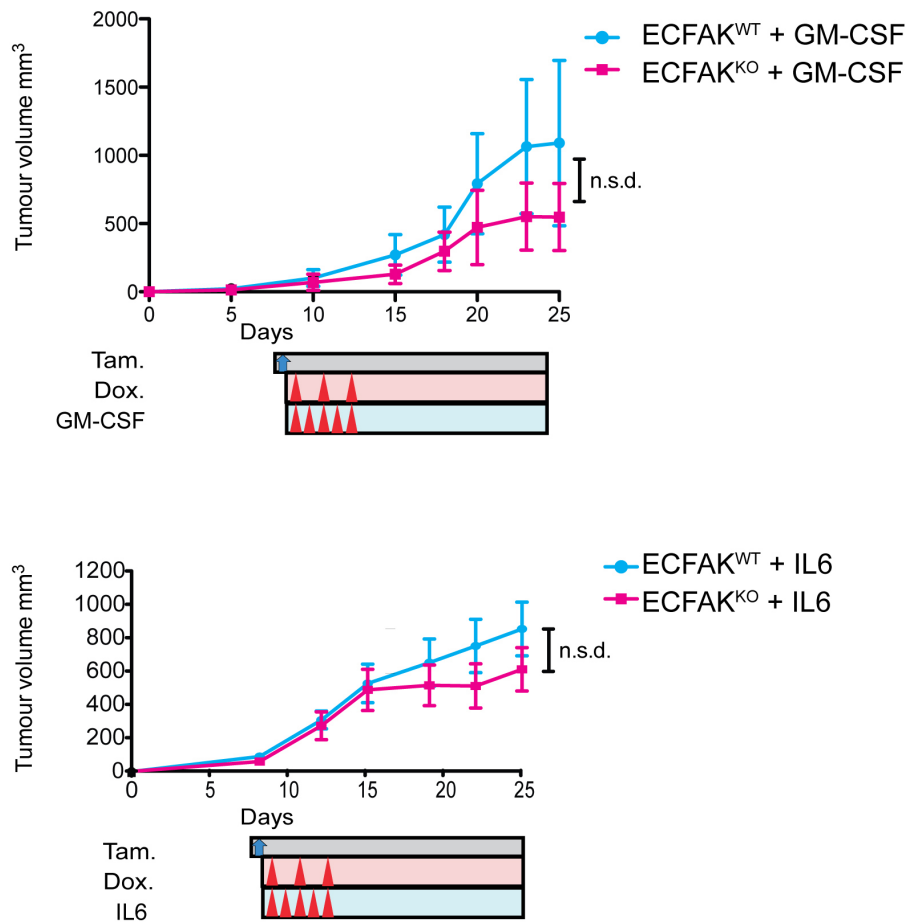
cytokine arrays. Bar chart shows the fold difference in cytokine expression between irradiated and non-irradiated wild-type and FAK-null endothelial cells + s.e.m.,  $n = 4$  experimental repeats. † $P = 0.06$ , \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , Student's  $t$ -test.



**Extended Data Figure 9 | Cytokine production is similar in untreated wild-type and FAK-null endothelial cells and DNA damage does not increase tumour-cell p-STAT3 expression in ECFAK<sup>KO</sup> mice.** **a**, Wild-type (WT) and FAK-null endothelial-cell whole-cell lysates were extracted and used in proteome profiler cytokine arrays. Bar chart shows baseline cytokine expression in untreated wild-type and FAK-null endothelial cells + s.e.m.,  $n = 4$  experimental repeats. NS, not significant. **b**, *Pdgfb-iCre<sup>ER</sup>;Fak<sup>fl/fl</sup>* mice and control mice (*Pdgfb-iCre<sup>ER</sup>*; non-floxed) were injected subcutaneously with CMT19T tumour cells (day 0). At 7–8 days post-inoculation, that is, once tumour growth was established, mice were given tamoxifen to induce

endothelial-cell FAK deletion in *Pdgfb-iCre<sup>ER</sup>;Fak<sup>fl/fl</sup>* but not *Pdgfb-iCre<sup>ER</sup>*; non-floxed mice, generating ECFAK<sup>KO</sup> and ECFAK<sup>WT</sup> mice, respectively. Thereafter CMT19T bearing mice were given 5 Gy gamma irradiation (day 10). Immunostaining for p-STAT3 in tumour sections revealed that although irradiation increased the percentage of p-STAT3-positive perivascular tumour cells this was not evident in ECFAK<sup>KO</sup> mice. Representative images of double immunostaining for p-STAT3 and PECAM are given. Bar chart shows mean percentage of p-STAT3-positive perivascular tumour cells + s.e.m.  $n = 6$  mice per group. \*\*\* $P < 0.005$ , Student's  $t$ -test. NS, not significant.





**Extended Data Figure 10 | In vivo rescue of chemosensitization phenotype.** Mouse melanoma B16F0 cells ( $1 \times 10^6$ ) were injected subcutaneously in the flank of *Pdgfb-iCre<sup>ER</sup>;Fak<sup>f/f</sup>* mice and control mice (*Pdgfb-iCre<sup>ER</sup>;non-floxed*). Seven days after tumour-cell injection, that is, once tumour growth was established, these mice were given tamoxifen to induce endothelial-cell FAK deletion in *Pdgfb-iCreER;Fak<sup>f/f</sup>* but not *Pdgfb-iCreER;non-floxed* mice,

generating ECFAC<sup>KO</sup> and ECFAC<sup>WT</sup> mice, respectively. Intratumoral injection of a low dose of recombinant GM-CSF (15 ng) (top graph), or IL-6 (3 ng) (bottom graph), restored doxorubicin-treated tumour growth in ECFAC<sup>KO</sup> mice to wild-type levels. Top graph shows mean tumour volumes over time + standard deviation. Bottom graph shows mean tumour volumes over time + s.e.m.  $n = 10-18$  mice per group. NS, not significant, Student's *t*-test.

# Promoter sequences direct cytoplasmic localization and translation of mRNAs during starvation in yeast

Brian M. Zid<sup>1,2</sup> & Erin K. O'Shea<sup>1,2,3,4</sup>

**A universal feature of the response to stress and nutrient limitation is transcriptional upregulation of genes that encode proteins important for survival. Under many such conditions, the overall protein synthesis level is reduced, thereby dampening the stress response at the level of protein expression<sup>1</sup>. For example, during glucose starvation in *Saccharomyces cerevisiae* (yeast), translation is rapidly repressed, yet the transcription of many stress- and glucose-repressed genes is increased<sup>2,3</sup>. Here we show, using ribosomal profiling and microscopy, that this transcriptionally upregulated gene set consists of two classes: one class produces messenger RNAs that are translated during glucose starvation and are diffusely localized in the cytoplasm, including many heat-shock protein mRNAs; and the other class produces mRNAs that are not efficiently translated during glucose starvation and are concentrated in foci that co-localize with P bodies and stress granules, a class that is enriched for mRNAs involved in glucose metabolism. Surprisingly, the information specifying the differential localization and protein production of these two classes of mRNA is encoded in the promoter sequence: promoter responsiveness to heat-shock factor 1 (Hsf1) specifies diffuse cytoplasmic localization and higher protein production on glucose starvation. Thus, promoter sequences can influence not only the levels of mRNAs but also the subcellular localization of mRNAs and the efficiency with which they are translated, enabling cells to tailor protein production to the environmental conditions.**

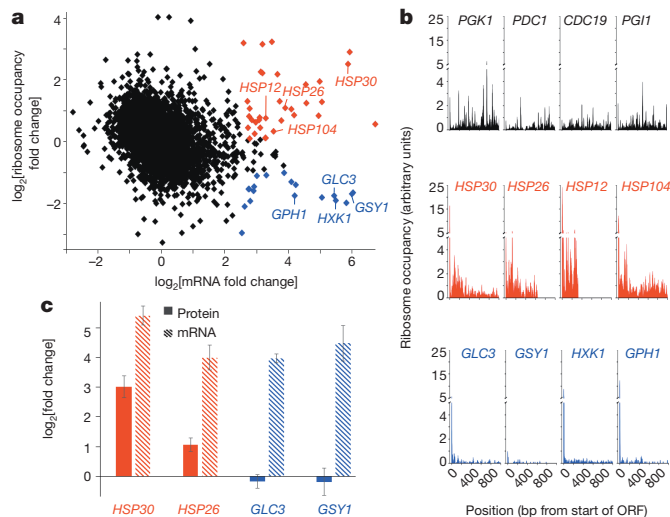
To investigate how cells alter gene expression during stress conditions that elicit an overall reduction in translation, we performed ribosomal profiling<sup>4</sup> on budding yeast cells grown in glucose-replete conditions and glucose-starvation conditions. In agreement with previous results<sup>2</sup>, during glucose starvation there was a collapse of polysomes into the 80S monosome peak, indicative of a reduction in global translation (Extended Data Fig. 1a). As reported previously<sup>3</sup>, we also observed an inverse correlation between the change in ribosome occupancy upon glucose starvation and the change in mRNA levels (Fig. 1a and Extended Data Fig. 2). For mRNAs whose levels increase during glucose starvation, we observed two classes of behaviour: some upregulated mRNAs ( $\log_2[\text{expression fold change}] > 2.5$ ) showed a decrease in ribosome occupancy upon glucose starvation ( $\log_2[\text{occupancy fold change}] < -1$ ; Fig. 1a, blue dots), whereas others showed an increase in ribosome occupancy that was greater than the median increase for all mRNAs ( $\log_2[\text{ribosome occupancy fold change}] > 0.09$ ; Fig. 1a, red versus black dots). Moreover, during glucose starvation we observed significantly higher ribosome occupancy in the coding region of the upregulated mRNAs with increased ribosome occupancy than that of the upregulated mRNAs with decreased ribosome occupancy (Fig. 1b; red versus blue genes). The upregulated mRNAs with higher ribosome occupancy were enriched in stress-response mRNAs (16 of 26 genes;  $P = 2.4 \times 10^{-9}$ ), including those encoding heat-shock proteins (Hsp) (Fig. 1a, b and Extended Data Table 1). By contrast, the upregulated mRNAs with lower ribosomal occupancy were enriched for those encoding proteins involved in glucose metabolism (7 of 18 genes;  $P = 7.8 \times 10^{-4}$ ).

Because there is a large reduction in global translation during glucose limitation, our measurements of ribosome occupancy under these conditions are almost certainly overestimates (see Methods). Although this overestimation increases the fold change in ribosome occupancy for all mRNAs, the relative differences between mRNAs are preserved (for example, red versus blue mRNAs). Moreover, we observed these relative differences in ribosome occupancy in different yeast strains and using different RNA isolation methods (Extended Data Fig. 2 and Supplementary Table 1). Thus, mRNAs that are upregulated during glucose starvation have differences in ribosome occupancy.

To determine whether the differences in ribosome occupancy translate into differences in protein production during glucose starvation, we measured protein levels by western blotting. We observed significant increases in proteins derived from the upregulated, higher ribosome occupancy mRNAs *HSP30* and *HSP26* (eightfold and twofold, respectively; Fig. 1c, red bars), but no significant change in protein levels was observed for proteins derived from the lower ribosome occupancy mRNAs *GLC1* and *GSY1* (Fig. 1c, blue bars), even though the mRNAs were induced to similar levels (Fig. 1c). For all upregulated genes that were assessed, we observed a corresponding increase in RNA polymerase II occupancy in their open reading frames (ORFs), suggesting that increased transcription contributes to the upregulation of the corresponding mRNAs during glucose starvation (Extended Data Fig. 3). Thus, upon glucose starvation, transcriptionally upregulated mRNAs have differences in ribosome occupancy, which lead to differences in protein production.

Since some mRNAs localize to messenger ribonucleoprotein (mRNP) foci (including P bodies and stress granules) during glucose limitation<sup>5</sup>, one possibility is that mRNA localization influences the ribosome occupancy and translational properties of an mRNA. To investigate whether mRNAs with differences in ribosome occupancy have differences in localization during glucose limitation, we generated fusions of gene coding regions with the MS2 sequence, and we visualized mRNAs using the MS2 coat protein (MS2-CP) fused to green fluorescent protein (GFP), and P bodies using red fluorescent protein (RFP) fused to the P body protein component Dcp2 (ref. 6). In agreement with previous observations<sup>7</sup>, *PGK1* and *PDC1* mRNAs, which are abundant pre-starvation, localized predominantly to P bodies after glucose starvation (Fig. 2a, b). By contrast, the transcriptionally upregulated, higher ribosome occupancy mRNAs *HSP26* and *HSP30* remained diffusely localized during glucose starvation, and the transcriptionally upregulated, lower ribosome occupancy mRNAs *GLC3* and *HXX1* became localized to P bodies as well as to other foci (Fig. 2a, b). The formation of foci was dependent on glucose starvation (Extended Data Fig. 4). Stress granules, which contain high concentrations of translation initiation factors, are formed during conditions in which translation is impaired and have been shown to partially overlap with P body foci<sup>8</sup>. Using a Pab1–cyan fluorescent protein (CFP) fusion to visualize stress granules<sup>9</sup>, we found that stress granules co-localize with a subset of P bodies, as well as with *GLC3* mRNA foci that were independent of P bodies (Fig. 2c). Therefore, mRNA classes with different

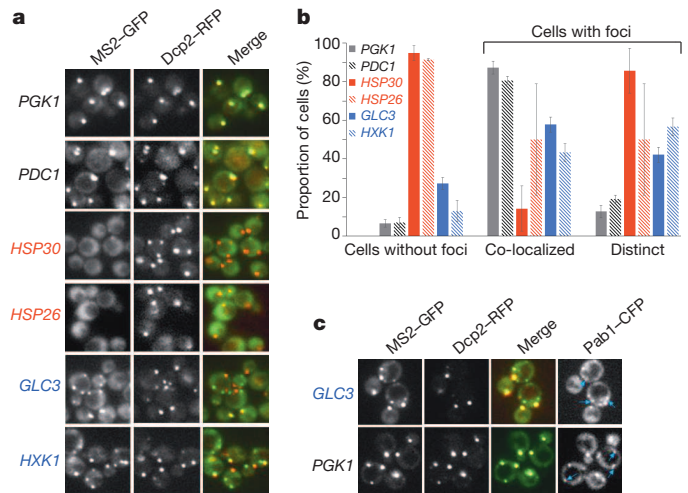
<sup>1</sup>Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>2</sup>Faculty of Arts and Sciences Center for Systems Biology, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>3</sup>Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA. <sup>4</sup>Howard Hughes Medical Institute, Harvard University, Cambridge, Massachusetts 02138, USA.



**Figure 1 | Ribosomal profiling reveals differences in ribosome occupancy of transcriptionally upregulated mRNAs upon glucose starvation.** **a**, Fold change in ribosome occupancy versus fold change in mRNA levels 15 min after cells are transferred to medium lacking glucose compared with levels in glucose-rich medium. mRNAs are represented by individual symbols on the plot. Ribosome occupancy was calculated for the coding region of each mRNA by dividing the total number of ribosome sequence counts in an ORF (normalized to the total number of aligned reads in reads per million reads (RPM)) by the number of mRNA sequence counts (RPM) in the same sequence. RNA sequencing was performed on RNA depleted of ribosomal RNA but not of poly(A)<sup>+</sup>-selected RNA. Red symbols denote genes that have upregulated mRNA levels ( $\log_2[\text{fold change}] > 2.5$ ) and a higher ribosome occupancy ( $\log_2[\text{fold change}] > 0.09$ ). Blue symbols denote genes that have upregulated mRNA levels ( $\log_2[\text{fold change}] > 2.5$ ) and a lower ribosome occupancy ( $\log_2[\text{fold change}] < -1.0$ ). Black symbols represent all other genes in the genome for which measurements were obtained. **b**, Ribosome occupancy (calculated as ribosome reads at each position relative to the average mRNA reads per base pair) for three classes of mRNA: those from genes that had high levels of the corresponding mRNA before glucose limitation (black); those whose mRNA levels increased during glucose limitation and had a higher ribosome occupancy during glucose limitation (red); and those whose mRNA levels increased during glucose limitation and had a lower ribosome occupancy (blue). The time point shown is 15 min after glucose starvation began. **c**, Strains expressing tandem affinity purification (TAP)-tagged versions of the indicated genes<sup>19</sup> grown in glucose-rich medium and then starved of glucose. The mRNA levels were measured by quantitative PCR after 15 min of glucose starvation. The protein abundance, as determined by western blotting, was measured after 30 min of glucose starvation. The mean fold changes in protein abundance (solid bars) and mRNA levels (striped bars)  $\pm$  s.e.m. were calculated relative to the respective values in glucose-rich medium. The western blotting experiments were performed on four independent biological replicates, and protein levels were normalized to Tub1 protein levels. The Hsp30 and Hsp26 protein levels were significantly higher upon glucose starvation than in glucose-rich medium ( $P < 0.05$ ). A one-tailed, paired  $t$ -test was used to determine  $P$  values. The mRNA measurements were made on three independent biological replicates and normalized to *ACT1* mRNA levels. bp, base pairs.

ribosome occupancy and protein production properties have distinct subcellular localization patterns.

To investigate whether the timing of mRNA production relative to glucose limitation contributes to mRNA localization, we analysed the localization and translation of a reporter gene that consisted of the doxycycline-inducible Tet-On promoter controlling the expression of *lacZ*-MS2. When *lacZ*-MS2 was induced before glucose starvation and cells were then starved, the mRNAs co-localized predominantly with P bodies (Extended Data Fig. 5a, b): this is consistent with the published observation that mRNAs that exist pre-starvation become localized to P bodies upon glucose limitation<sup>7</sup>. By contrast, when the mRNA was induced only during glucose starvation, it formed foci that co-localized with P

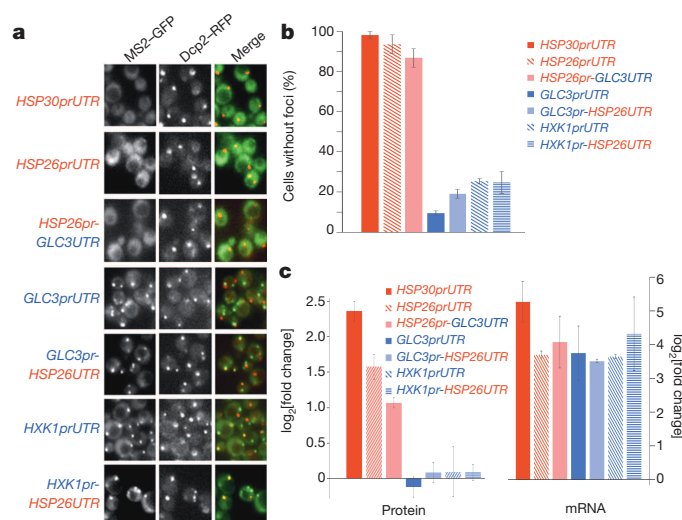


**Figure 2 | Glucose starvation induces differences in the localization of mRNAs.** **a**, The promoter, 5' UTR and ORF of genes of interest were fused upstream of the MS2 sequence and the native 3' UTR, and then mRNA was visualized with fluorescence microscopy using an MS2-coat-protein-GFP fusion reporter after 15 min of glucose starvation. Dcp2-RFP was used to visualize P bodies<sup>6</sup>. **b**, Quantification of mRNA localization to P bodies, as measured by overlap of the MS2 signal with that of the P body marker Dcp2-RFP. Cells without foci were cells that had P body foci but no GFP foci. For cells that contained both GFP (MS2) and RFP (P body) foci, the foci were categorized as either co-localized with P bodies or not overlapping with (distinct from) P bodies. The values are presented as mean  $\pm$  s.e.m. of the data from Fig. 2a, measured on a minimum of 25 cells in quadruplicate (two biological replicates with two technical replicates per sample). Cells expressing *HSP30* or *HSP26* mRNA had fewer foci than cells expressing *PGK1*, *PDC1*, *GLC3* or *HXK1* mRNAs ( $P < 0.01$ ). Cells expressing *GLC3* or *HXK1* mRNA had significantly more distinct foci than those expressing *PGK1* or *PDC1* mRNA ( $P < 0.05$  for all comparisons). Since few cells expressing *HSP30* or *HSP26* mRNA had foci, these cells were excluded from statistical analyses because of the high measurement variability. A two-tailed, two-sample unequal variance  $t$ -test was performed to determine the  $P$  values. **c**, Pab1-CFP was used to visualize stress granule localization<sup>10</sup> after 30 min of glucose starvation. The blue arrows point to stress granules (Pab1-CFP signals) that do not co-localize with P bodies (Dcp2-RFP signal). In the *GLC3*-MS2 strain, the Pab1-CFP stress granule foci overlapped with the GFP foci that were distinct from P bodies.

bodies and foci distinct from P bodies, a pattern similar to that of transcriptionally upregulated, lower ribosome occupancy mRNAs (Fig. 2, blue mRNAs). These results were not sensitive to the timing of induction during starvation or to the level of induction (over a range from 4-fold to 30-fold induction) (Extended Data Fig. 5). Thus, the timing of mRNA production relative to glucose limitation influences cytoplasmic mRNA localization.

The timing of mRNA production can influence whether mRNAs are localized exclusively to P bodies, but it is unclear what causes the differential localization and translation of transcriptionally upregulated, higher and lower ribosome occupancy mRNAs. To determine whether we could identify signals present in the mRNA itself, we fused the promoter and/or 5' untranslated region (UTR) of each gene to a constant ORF, *CFP*, and found that these fusions exhibited the same patterns of localization as the native ORFs, suggesting that the information specifying localization was contained in these elements (Fig. 3a, b and Extended Data Fig. 6a). To determine whether the promoter or 5' UTR was sufficient to determine localization, we generated chimaeras between the *HSP26* promoter and the *GLC3* 5' UTR, as well as between the *GLC3* or *HXK1* promoter and the *HSP26* 5' UTR. In each case the promoter was sufficient to recapitulate the localization observed for the native gene (Fig. 3a, b and Extended Data Fig. 6a). Changes in the transcription start sites are not likely to account for these observations, as we did not observe significant differences in the 5' ends of the mRNAs produced from the chimaeras





**Figure 3 | Differential localization of mRNAs is determined by the promoter.** **a**, The promoter (pr) and 5' UTR of the indicated genes were fused upstream of *CFP-MS2*, and localization of the resultant mRNA was visualized 15 min after glucose starvation. **b**, Quantification of the data shown in Fig. 3a. The values are presented as the mean  $\pm$  s.e.m. measured on a minimum of 30 cells in quadruplicate (two biological replicates with two technical replicates per sample). There were significantly fewer foci in the cells containing *HSP30*- or *HSP26*-promoted mRNAs (red bars) than in cells with *GLC3*- or *HXK1*-promoted mRNAs (blue bars) ( $P < 0.01$ ). A two-tailed, two-sample unequal variance *t*-test was performed to determine the *P* values. **c**, Protein expression directed by the indicated promoter-UTR combinations was measured by western blotting using an anti-GFP antibody that recognizes CFP. The protein levels were measured after 30 min of glucose starvation and in glucose-rich medium, and the fold change in protein abundance was calculated. *MS2-CP-GFP(3 $\times$ )* driven by the *MYO2* promoter was used as a loading control for western blotting. The protein fold changes are presented as mean  $\pm$  s.e.m. and were calculated from four independent biological replicates. The levels of the proteins produced from *HSP30*- and *HSP26*-promoted mRNAs (red bars) significantly increased upon glucose starvation compared with in glucose-rich medium ( $P < 0.01$ ). A one-tailed, paired *t*-test was used to determine the *P* values. The fold change in mRNA levels was determined after 15 min of glucose starvation versus growth in glucose-rich medium, measured by qPCR using *CFP* primers and *ACT1* levels to normalize expression. The mRNA fold-change values are presented as mean  $\pm$  s.e.m. and were calculated from three independent biological replicates.

(Extended Data Table 2). To determine whether the correlation between localization and translation that we observed previously for native mRNAs (Figs 1 and 2) also holds for these chimaeras, we measured protein production and found that the *HSP* promoters that specify diffuse mRNA localization result in a larger increase in protein production during glucose starvation (Fig. 3c, red bars). By contrast, although the focus-forming *GLC3* and *HXK1* promoters drive levels of mRNA induction similar to those driven by the *HSP26* promoter, during glucose starvation there was no significant increase in protein production from mRNAs driven by these promoters (Fig. 3c, blue bars). Thus, promoters can influence gene expression through means other than by simply controlling mRNA induction.

To identify specific promoter sequences that influence mRNA localization and protein production, we made promoter chimaeras from *GLC3* and *HSP26* and used these chimaeras to drive the expression of *CFP-MS2*. Two sets of transcription factors that activate transcription upon glucose starvation are Msn2 and Msn4, which bind to stress-response elements (STREs)<sup>10</sup>, and Hsf1, which binds to heat-shock elements (HSEs)<sup>11</sup>. We generated chimaeric promoters containing combinations of STREs and HSEs from the *GLC3* and *HSP26* promoters and analysed mRNA localization of mRNAs translated from reporter constructs. We found that mRNA reporters whose expression was controlled by chimaera 1 or 4 were induced in response to glucose limitation, with the mRNA

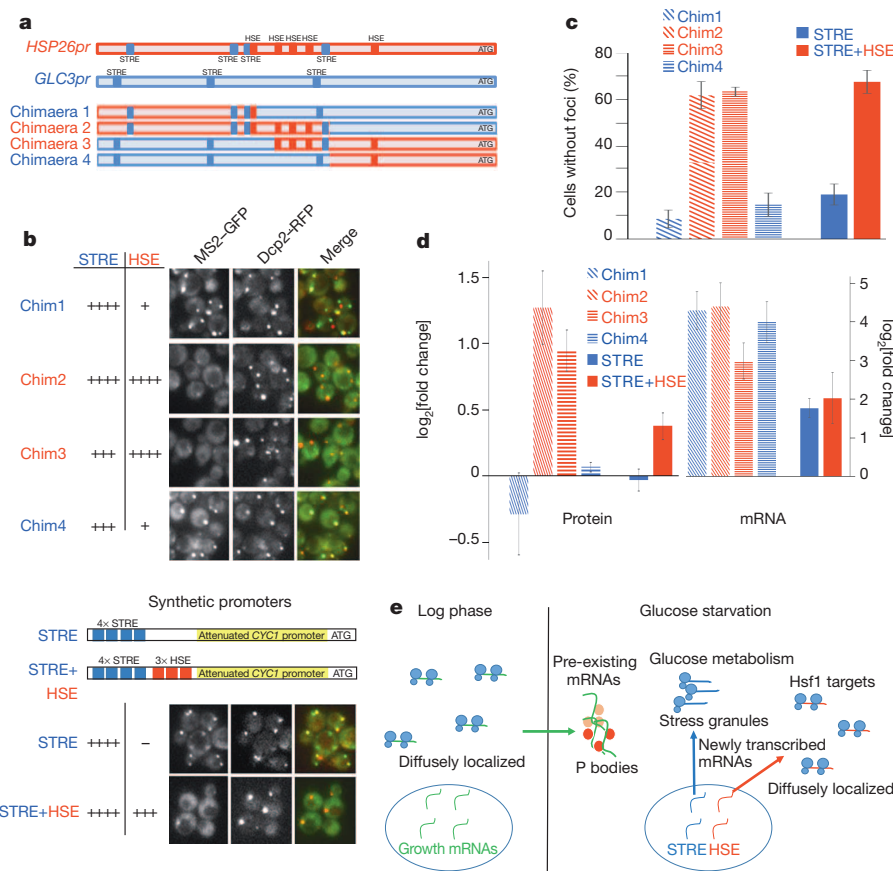
forming foci in a high percentage of cells, but they produced no significant change in CFP protein levels (Fig. 4b–d and Extended Data Fig. 6b): these chimaeras exclude many of the HSE sites contained in the *HSP26* promoter but include at least three STRE sites. By contrast, mRNA reporters whose expression was controlled by chimaera 2 or 3 were also induced in response to glucose limitation, but the mRNA had generally diffuse localization (Fig. 4b–d and Extended Data Fig. 6b), more similar to that of the full-length *HSP26* promoter; they also produced significant increases in protein levels upon glucose starvation (Fig. 4d). All four chimaeras resulted in similar levels of mRNA induction, and chimaeras 1 and 2 led to similar absolute mRNA levels after 15 min of glucose starvation (Fig. 4d and Extended Data Fig. 7a). The sequence in common to chimaeras 2 and 3 is a 90-base-pair region that contains several HSEs<sup>12</sup>.

To determine whether Hsf1 responsiveness correlates with the diffuse localization of the chimaeras, we treated cells expressing different reporters with azetidine-2-carboxylate (AZC), a proline analogue that robustly increases *Hsf1* transcriptional activity with low activation of the STRE response<sup>13</sup>. The full-length *HSP26* and *GLC3* promoters both exhibited strong induction in response to glucose starvation, but only the *HSP26* promoter showed robust induction in response to AZC treatment (Extended Data Fig. 7b). All four of the chimaeric promoters responded similarly to glucose starvation, but only chimaeras 2 and 3 showed greater than 20-fold induction upon treatment with AZC (Extended Data Fig. 7b). Thus, responsiveness to Hsf1 correlates with, and may contribute to, diffuse mRNA localization and protein production during glucose starvation.

Our ability to assess whether Hsf1 is necessary for diffuse localization was precluded by the *S. cerevisiae Hsf1* deletion strain being non-viable<sup>14</sup>; however, we can infer the role of Hsf1 by examining the localization of mRNAs and the protein production resulting from constructs that contain different combinations of STREs and HSEs (Fig. 4b, schematic, and Extended Data Fig. 8). *CFP-MS2* expressed under the control of a synthetic promoter that contains only STREs formed many foci during glucose limitation (Fig. 4b, bottom panel, Fig. 4c, right, and Extended Data Fig. 6b). The addition of three HSEs to this synthetic STRE promoter was sufficient to switch the mRNA localization from foci to diffuse localization (Fig. 4b, bottom panel, Fig. 4c, right, and Extended Data Fig. 6b). The synthetic reporter containing HSE binding sites resulted in more protein production during glucose starvation than a synthetic promoter containing only STREs, even though the two promoters had similar levels of mRNA induction (Fig. 4d). We conclude that HSE binding sites, probably functioning through the transcription factor Hsf1, influence mRNA localization and translation upon glucose starvation.

Our data suggest that promoter sequences and the action of select transcription factors in the nucleus can influence mRNA localization and translation upon glucose starvation (Fig. 4e). A link between transcriptional regulation and cytoplasmic localization may be a general adaptation during times of stress, enabling cells to coordinately regulate the production of entire classes of proteins. Under non-stress conditions, the upregulation of a class of transcripts by a transcription factor would produce similar amounts of protein from each of the mRNAs, as translation would proceed at a generally high rate. Under stressful conditions, when translation is reduced overall, selective translation may be required to produce proteins that are needed for adaptation to the new conditions. In the case of glucose starvation, Hsf1 targets that encode cytoprotective proteins, such as chaperones, may need to be produced as soon as possible to help the cells cope with the stress, but alternative glucose metabolism genes may be superfluous when no carbon source is present. The induction of mRNA without a concomitant increase in the protein level of genes involved in glucose metabolism may allow cells to more rapidly produce proteins upon reintroduction of a carbon source. Intriguingly, the localization of *HSP70* and *HSP90* mRNA in stressed yeast and mammalian cells appears to be similar: these mRNAs are largely excluded from stress granules during cellular stress in mammalian cells<sup>15,16</sup>. Previous studies have shown that the promoter can influence the stability of an mRNA through co-transcriptional loading of an





**Figure 4 | Promoter elements influence mRNA translation efficiency and localization.** **a**, Schematic of the promoter regions that control the expression of CFP-MS2. The HSE sites are taken from ref. 12. **b**, Localization of CFP-MS2 mRNA driven by chimaeric (Chim) HSP26 and GLC3 promoters or synthetic promoters (STRE or STRE+HSE) visualized after 15 min of glucose starvation. Dcp2-RFP was used to visualize P bodies. **c**, Quantification of the data shown in Fig. 4b. The values are presented as the mean  $\pm$  s.e.m. and were measured on a minimum of 25 cells in quadruplicate (two biological replicates with two technical replicates per sample). Chim1- or Chim4-promoted mRNAs showed more focus formation than did Chim2- or Chim3-promoted mRNAs ( $P < 0.05$ ). The promoter containing only STRE (and no HSE) resulted in mRNA showing more focus formation than did the STRE+HSE-promoted mRNA ( $P < 0.01$ ). A two-tailed, two-sample unequal variance  $t$ -test was performed to determine the  $P$  values. **d**, Protein expression directed by the different promoters was measured by western blotting using an anti-GFP antibody that recognizes CFP. The protein levels were measured after 30 min of glucose starvation, and the fold change was calculated relative to measurements in glucose-rich medium. MS2-CP-GFP(3 $\times$ ) driven by the MYO2 promoter was used as a loading control for western blotting. The protein

fold-change values are presented as the mean  $\pm$  s.e.m. and were calculated from five independent biological replicates for the chimaeric promoters and seven independent biological replicates for the synthetic promoters. The levels of proteins produced from Chim2-, Chim3- and STRE+HSE-driven mRNAs (red bars) were significantly increased upon glucose starvation compared with in glucose-rich medium ( $P < 0.05$ ). A one-tailed, paired  $t$ -test was used to determine the  $P$  values. The fold change in mRNA levels was measured after 15 min of glucose starvation versus growth in glucose-rich medium, by using qPCR with CFP primers and with expression normalized to ACT1 levels. The mRNA fold-change values are presented as the mean  $\pm$  s.e.m. and were calculated from three independent biological replicates. **e**, mRNAs that are actively transcribed during the logarithmic (log) phase, such as PDC1 and PGK1, localize to P bodies during glucose starvation. mRNAs whose expression is activated by Hsf1 through HSE elements, including HSP30 and HSP26, are diffusely localized during glucose starvation and result in high levels of protein production. mRNAs such as GLC3 and HXX1, which are upregulated by STREs independently of Hsf1, localize to P bodies and stress granules and are poorly translated during glucose starvation.

accessory protein onto the mRNA<sup>17,18</sup>. A similar phenomenon may be operating here, whereby transcription factors load RNA-binding proteins that direct mRNA localization, or there may be *cis* alterations to the mRNA, such as to its poly(A)<sup>+</sup> tail length or methylation, that influence its fate. Future studies will reveal aspects of these mechanisms, as well as whether this phenomenon is conserved in other eukaryotes.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 October 2013; accepted 11 June 2014.

Published online 3 August 2014.

1. Simpson, C. E. & Ashe, M. P. Adaptation to stress in yeast: to translate or not? *Biochem. Soc. Trans.* **40**, 794–799 (2012).
2. Ashe, M. P., De Long, S. K. & Sachs, A. B. Glucose depletion rapidly inhibits translation initiation in yeast. *Mol. Biol. Cell* **11**, 833–848 (2000).

3. Arribere, J. A., Doudna, J. A. & Gilbert, W. V. Reconsidering movement of eukaryotic mRNAs between polysomes and P bodies. *Mol. Cell* **44**, 745–758 (2011).
4. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S. & Weissman, J. S. Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223 (2009).
5. Decker, C. J. & Parker, R. P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. *Cold Spring Harb. Perspect. Biol.* **4**, a012286 (2012).
6. Brengues, M., Teixeira, D. & Parker, R. Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. *Science* **310**, 486–489 (2005).
7. Teixeira, D., Sheth, U., Valencia-Sanchez, M. A., Brengues, M. & Parker, R. Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA* **11**, 371–382 (2005).
8. Buchan, J. R. & Parker, R. Eukaryotic stress granules: the ins and outs of translation. *Mol. Cell* **36**, 932–941 (2009).
9. Hoyle, N. P., Castelli, L. M., Campbell, S. G., Holmes, L. E. & Ashe, M. P. Stress-dependent relocalization of translationally primed mRNPs to cytoplasmic granules that are kinetically and spatially distinct from P-bodies. *J. Cell Biol.* **179**, 65–74 (2007).

10. Martínez-Pastor, M. T. *et al.* The *Saccharomyces cerevisiae* zinc finger proteins Msn2p and Msn4p are required for transcriptional induction through the stress response element (STRE). *EMBO J.* **15**, 2227–2235 (1996).
11. Hahn, J.-S. & Thiele, D. J. Activation of the *Saccharomyces cerevisiae* heat shock transcription factor under glucose starvation conditions by Snf1 protein kinase. *J. Biol. Chem.* **279**, 5169–5176 (2004).
12. Susek, R. E. & Lindquist, S. Transcriptional derepression of the *Saccharomyces cerevisiae* HSP26 gene during heat shock. *Mol. Cell. Biol.* **10**, 6362–6373 (1990).
13. Trotter, E. W. *et al.* Misfolded proteins are competent to mediate a subset of the responses to heat shock in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **277**, 44817–44825 (2002).
14. Wiederrecht, G., Seto, D. & Parker, C. S. Isolation of the gene encoding the *S. cerevisiae* heat shock transcription factor. *Cell* **54**, 841–853 (1988).
15. Kedersha, N. & Anderson, P. Stress granules: sites of mRNA triage that regulate mRNA stability. *Biochem Soc. Trans.* **30**, 963–969 (2002).
16. Stöhr, N. *et al.* ZBP1 regulates mRNA stability during cellular stress. *J. Cell Biol.* **175**, 527–534 (2006).
17. Bregman, A. *et al.* Promoter elements regulate cytoplasmic mRNA decay. *Cell* **147**, 1473–1483 (2011).
18. Trcek, T., Larson, D. R., Moldón, A., Query, C. C. & Singer, R. H. Single-molecule mRNA decay measurements reveal promoter-regulated mRNA stability in yeast. *Cell* **147**, 1484–1497 (2011).
19. Ghaemmaghami, S. *et al.* Global analysis of protein expression in yeast. *Nature* **425**, 737–741 (2003).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** We thank A. Subramaniam, S. Mukherji, V. Denic, B. Stern and M. Waldram for feedback on the paper, members of the O'Shea, Denic, and Calarco laboratories for discussions, and C. Daly for sequencing assistance. This research was supported by the American Cancer Society and the New England Division Funding A Cure initiative (B.M.Z.) and the Howard Hughes Medical Institute (E.K.O.).

**Author Contributions** B.M.Z. collected and analysed the data. B.M.Z. and E.K.O. designed the experiments and wrote the manuscript.

**Author Information** All raw sequencing data have been deposited in the Gene Expression Omnibus database under accession number GSE56622. Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to E.K.O. (Erin\_Oshea@harvard.edu).

## METHODS

**Yeast strains and growth.** All yeast strains are listed in Supplementary Table 2. For the ribosomal profiling experiments presented in Fig. 1, the yeast strain BY4741 (Euroscarf) (MATa *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*) was grown at 30 °C in batch culture with shaking at 125 r.p.m. in synthetic complete glucose medium (SCD medium) and synthetic complete medium lacking glucose (SC medium). Yeast cells were randomly chosen by taking half of the cells from a culture for glucose starvation and the other half for assessment under glucose-replete conditions. There was no blinding to the group that the yeast were allocated to. Ribosomal profiling and RNA sequencing (RNA-seq) were repeated under the same growth conditions using the yeast strain EY0690 (W303 MATa *trp1-1 leu2-3 ura3-1 his3-11 can1-100*), and similar results were obtained (Extended Data Fig. 2 and Supplementary Table 1).

The yeast strain background W303 (EY0690) was used for all microscopy experiments. To image mRNAs, the 12× MS2 sequence was excised from the MS2L construct<sup>20</sup> and placed in the *LEU2*-marked integration vector pRS305. The *ADH1* 3' UTR was cloned downstream of this sequence, and gene-specific sequences were cloned upstream of the MS2 sequence by Gibson Assembly<sup>21</sup>. Integration was performed by cleaving the resultant plasmids within the *LEU2* gene with EcoRV and transforming the linear fragment into a yeast strain (EY2888) containing *MS2-CP-GFP*(3×)<sup>20</sup> under the control of the *MYO2* promoter integrated at the *HIS3* locus. Dcp2 was tagged with RFP, and Pab1 with the CFP mTurquoise2 by carboxy-terminal chromosomal integration of PCR products, including auxotrophic or antibiotic markers flanked by 40 base pairs (bp) of sequence found directly upstream and downstream of the gene, followed by selection on the appropriate medium<sup>22</sup>. *TET07-lacZ-MS2-ADH1\_3' UTR* was integrated into the EY2888 strain at the *LEU2* locus with the addition of the rtTA activator (Tet-On) under the control of the *ERV14* promoter (EB1674). Doxycycline was added to a final concentration of 20 µg ml<sup>-1</sup> either 15 min before glucose starvation or at the time of glucose starvation. A yeast-codon-optimized CFP reporter (*SCFP3A*)<sup>23</sup> was used to generate a uniform ORF for localization constructs driven by various promoters (Figs 3 and 4). To determine Hsf1 responsiveness, AZC was added to a final concentration of 10 mM to cells at an optical density at 600 nm (OD<sub>600</sub>) of ~0.2, and cells were incubated with shaking for 2 h.

Chimaeric *HSP26* and *GLC3* promoters, shown in Fig. 4a, were made by fusion PCR of the indicated promoter regions (Supplementary Table 2) and were inserted upstream of *CFP-MS2-ADH1\_3' UTR*. Synthetic reporters were created by Gibson Assembly of a 4× STRE with or without a 3× HSE element followed by an attenuated *CYC1* promoter<sup>24</sup> upstream of *CFP-MS2-ADH1\_3' UTR* (Extended Data Fig. 8). A Mig1-binding site<sup>25</sup> (AAAAATGCGGGG) was included 5' of the STREs to reduce leaky expression under glucose-rich conditions.

**Ribosomal profiling and RNA-seq.** Ribosomal profiling and RNA-seq were performed as described previously<sup>4</sup>. We performed two ribosomal profiling experiments for BY4741 and two for EY0690. Yeast were grown in SCD at 30 °C to an OD<sub>600</sub> between 0.3 and 0.4. Then, cells were collected by filtration, resuspended in SC medium (lacking glucose) and grown for 15 min. Cycloheximide was added to a final concentration of 100 µg ml<sup>-1</sup> for 1 min with continued shaking at 30 °C, and cells were then harvested. Cells were pulverized in a PM 100 ball mill (Retsch), and extracts were digested with RNase I followed by the isolation of ribosome-protected fragments either by purifying RNA from the monosome fraction of a sucrose gradient (BY4741 two samples) or by using a sucrose cushion (EY0690 two samples). Isolated 28-base sequences were polyadenylated, and reverse transcription was performed using either OTi225 (BY4741) or OTi9pA (EY0690) (Supplementary Table 3). OTi9pA allowed samples to be multiplexed at subsequent steps. RNA-seq was performed on RNA depleted of rRNA using a yeast Ribo-Zero kit (Epicentre) (EY0690, one experiment), total RNA (BY4741 and EY0690, one experiment each) or poly(A)<sup>+</sup>-selected RNA using Oligo(dT) Dynabeads (Invitrogen) (BY4741 and EY0690, two independent samples each). rRNA-depleted RNA and total RNA from EY0690 had a high Pearson correlation between samples ( $r > 0.9$ ), so these sequences were combined to give higher sequence coverage for the mRNA sample. BY4741 samples were sequenced on an Genome Analyzer II (Illumina), while EY0690 samples were multiplexed and sequenced on a HiSeq analyser (Illumina) (both at the FAS Center for Systems Biology Core Facility). All raw sequencing data are available at NCBI GEO, with accession number GSE56622.

To analyse the ribosomal profiling and RNA-seq sequences, reads were trimmed of the 3' run of poly(A)s and then aligned against *S. cerevisiae* rRNA sequences using Bowtie sequence aligner<sup>26</sup>. Reads that did not align to rRNA sequences were aligned against the full *S. cerevisiae* genome. Reads that had an unambiguous alignment with less than three mismatches were used in the measurements of ribosome occupancy and mRNA levels. Since there were many reads mapping to the initiation region (−16 bp to +20 bp in relation to the AUG; Extended Data Fig. 1b), the ribosome occupancy for each mRNA was calculated by taking the total number of ribosome reads (normalized to the total number of aligned reads in reads per million reads (RPM)) in the downstream region (+20 bp from the AUG to the end of the

ORF; Extended Data Fig. 1b) and dividing this by the number of mRNA reads (RPM) in the same region. The ribosome occupancy along the mRNA (Fig. 1b) was calculated by dividing the ribosome read counts at each base pair along the gene by the average number of mRNA reads per base pair for each gene. Because there is a large reduction in global translation during glucose limitation, our measurements of ribosome occupancy under these conditions are almost certainly overestimates. This arises because even when there is a large reduction in ribosomes associated with mRNAs, as seen by the collapse in the polysome profile during glucose starvation (Extended Data Fig. 1), we isolated and sequenced the same number of ribosome-protected sequence reads. Although this has the effect of increasing ribosome occupancy values for all genes, the relative differences between mRNAs remain (for example, red (*HSP* mRNAs) versus blue (glucose metabolism mRNAs)).

To reduce sampling error, a cut-off of 30 or more reads in the downstream region was set for RNA-seq during glucose starvation. Since we focused on mRNAs that were upregulated during glucose limitation and since many of these genes are poorly expressed in glucose-rich conditions, we set a cut-off of four or more reads for RNA-seq during glucose-rich conditions, as well as four or more reads for ribosomal profiling in both glucose-rich and glucose-starvation conditions. Even with such a low number of reads as a cut-off, there was a large overlap between ribosomal profiling and non-poly(A)<sup>+</sup>-selected RNA-seq experiments performed in BY4741 and EY0690 at both the individual gene level and the gene ontology class level for the different categories of upregulated mRNAs (Extended Data Table 1 and Supplementary Table 1). At the individual gene level, 21 of the 33 upregulated, higher ribosome occupancy genes classified from BY4741 were in the same category for the EY0690 data, while only 2 genes switched to the upregulated, lower ribosome occupancy category (Supplementary Table 1). For lower ribosome occupancy mRNAs, 13 of the 19 genes classified as such in BY4741 were also found in the same category for EY0690, while only 2 genes switched to the upregulated, higher ribosome occupancy category (Supplementary Table 1). Extended Data Fig. 2 shows the data from all four ribosomal profiling data sets, together with the six mRNA preparations that include both poly(A)<sup>+</sup>-selected and non-poly(A)<sup>+</sup>-selected mRNA.

**Polysome analysis.** Sucrose density gradients (10–50%) were prepared and measured using a BioComp Gradient Station (BioComp Instruments) according to the manufacturer's instructions. Sucrose solutions were prepared in 20 mM Tris, pH 8.0, 140 mM KCl, 5 mM MgCl<sub>2</sub>, 0.5 mM dithiothreitol (DTT) and 50 U ml<sup>-1</sup> SUPERase<sup>™</sup> In (Ambion). Samples were loaded onto gradients and spun for 3 h at 35,000 r.p.m. at 4 °C in an SW41 rotor (Beckman Coulter). Samples were passed through a BioComp Gradient Station, and the absorbance at 260 nm was read using an Econo UV Monitor (Bio-Rad).

**Microscopy.** Cells were grown to an OD<sub>600</sub> between 0.3 and 0.4 in SCD at 30 °C and then washed and resuspended in SC. After 15 min, cells were concentrated and imaged using an Axiovert 200M inverted microscope (Zeiss) with a Cascade 512 cooled charge-coupled (CCD) camera (Photometrics) and an oil-immersion 63× objective. A custom MATLAB script was written to measure co-localization of GFP mRNA foci and RFP P body foci. In brief, a threshold mask was set for individual cells using the Otsu Thresholding Filter<sup>27</sup> and subsequently used to create a binary image. The centroid of each focus was then obtained using the regionprops command. If no mRNA foci were found, the cell was counted as without foci. If there were one or more mRNA foci, the minimum distance between each mRNA focus and every P body foci in the cell was calculated. If this distance was less than or equal to 1.5 pixels, the focus was considered to be co-localized; otherwise, it was considered non-overlapping and distinct. *P* values were calculated using a two-tailed, two-sample unequal variance *t*-test to account for possible differences in variance, which may arise from unrelated data<sup>28</sup>. For stress granule visualization, cells were imaged after 30 min of glucose starvation to observe clear stress granule formation.

**Quantitative PCR (qPCR).** RNA was extracted using the MasterPure Yeast RNA Purification Kit (Epicentre). cDNA was prepared using SuperScript III Reverse Transcriptase (Invitrogen) with a combination of oligo(dT) primers and random hexamers according to the manufacturer's instructions. mRNA abundance was determined by qPCR using SYBR Green PCR Master Mix (Applied Biosystems) and primers specific for each transcript (Supplementary Table 3). The mRNA levels were normalized to *ACT1* abundance, and the fold change between glucose-starvation and glucose-rich conditions was calculated.

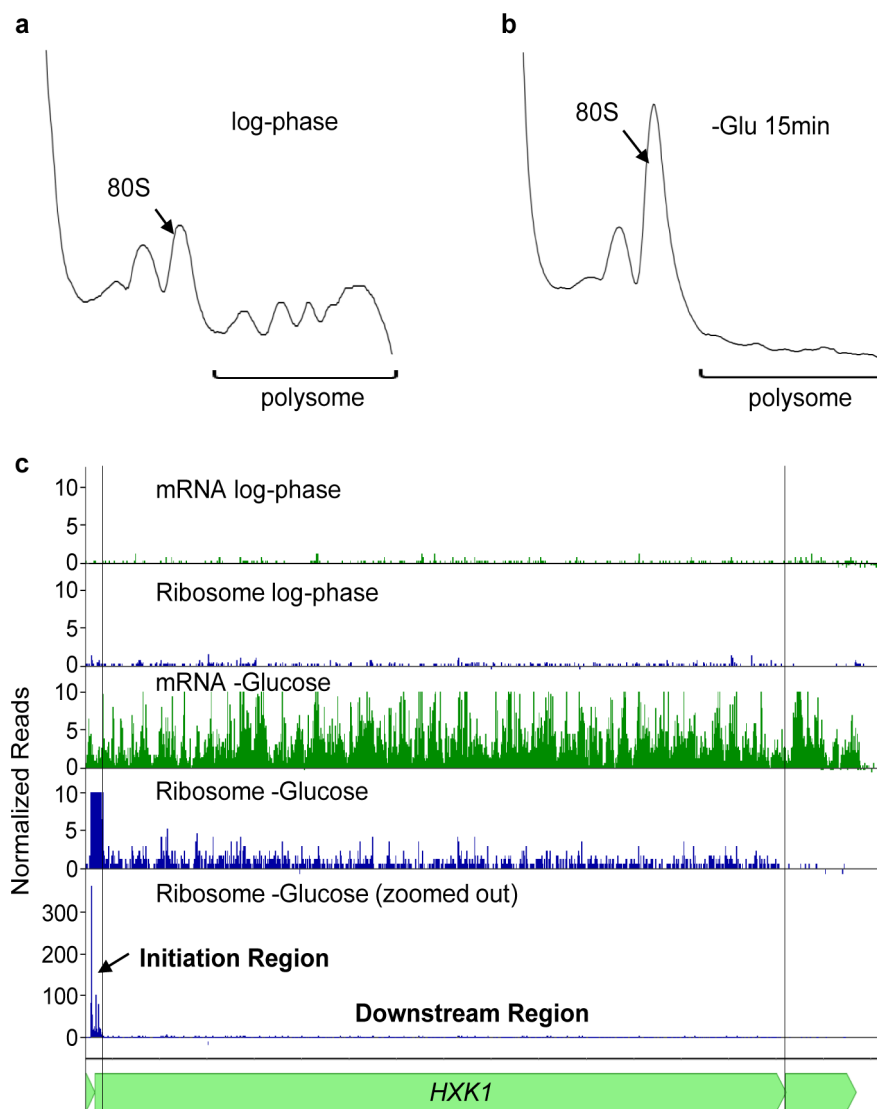
**Chromatin immunoprecipitation (ChIP).** ChIP–qPCR experiments were conducted as previously described<sup>29</sup> with the following differences. Rpb3–TAP (tandem affinity purification)<sup>19</sup> was used to determine RNA polymerase II occupancy in glucose-rich conditions and after 15 min of glucose starvation. Rpb3–TAP was immunoprecipitated using IgG Sepharose Fast Flow (GE Healthcare). Input and immunoprecipitated samples were assayed by qPCR to assess the extent of RNA polymerase II occupancy in different genomic regions. Primer pairs against the indicated ORFs, as well as an untranscribed telomeric region (Supplementary Table 3), were used to determine PCR efficiencies during glucose-rich and glucose-starvation conditions.

**Western blotting.** Strains were grown in the appropriate medium and then centrifuged at 4,000g for 2 min. Pellets were resuspended in Buffer A (0.5% Triton X-100, 150 mM NaCl, 1 mM EDTA and 50 mM HEPES, pH 7.4) followed by lysis with glass beads at 4 °C and centrifugation at 5,000g for 5 min. The crude extract was then resolved by SDS-PAGE, and a rabbit polyclonal antibody specific for calmodulin-binding peptide (A00635-40, GenScript) was used to detect TAP-tagged proteins. A mouse anti- $\alpha$ -tubulin antibody (12G10, Developmental Studies Hybridoma Bank) was used as a loading control. CFP and GFP were detected using a rabbit polyclonal antibody against GFP (A-6455, Invitrogen), with the *pMYO2*-driven *MS2-CP-GFP*(3 $\times$ ) used as a loading control. To determine whether there was an increase in protein levels from glucose-replete conditions to glucose-starvation conditions, a one-tailed, paired *t*-test was used. The Shapiro–Wilk statistic was computed to test for normality in these small sample sizes of four to seven replicates. These sample sizes are commonly used to measure differences in protein levels.

**5' RACE.** The transcriptional start site was determined for various promoters using the ExactSTART Eukaryotic mRNA 5' RACE Kit (Epicentre). An adaptor oligonucleotide (5' adaptor) was ligated to the 5' end of the RNA, and cDNA was synthesized using an oligo(dT) primer that contained another adaptor sequence (3' adaptor). The 5' region of the mRNA was amplified by PCR using a kit-provided 5' adaptor primer (5'-TCATACACATACGATTAGGTGACACTATAGAGCGCCGCCTGCAGGAAA-3'), and a CFP-specific primer (Supplementary Table 3). The PCR products were cloned into the pCR4-TOPO vector (Invitrogen) and were sequenced (Eton).

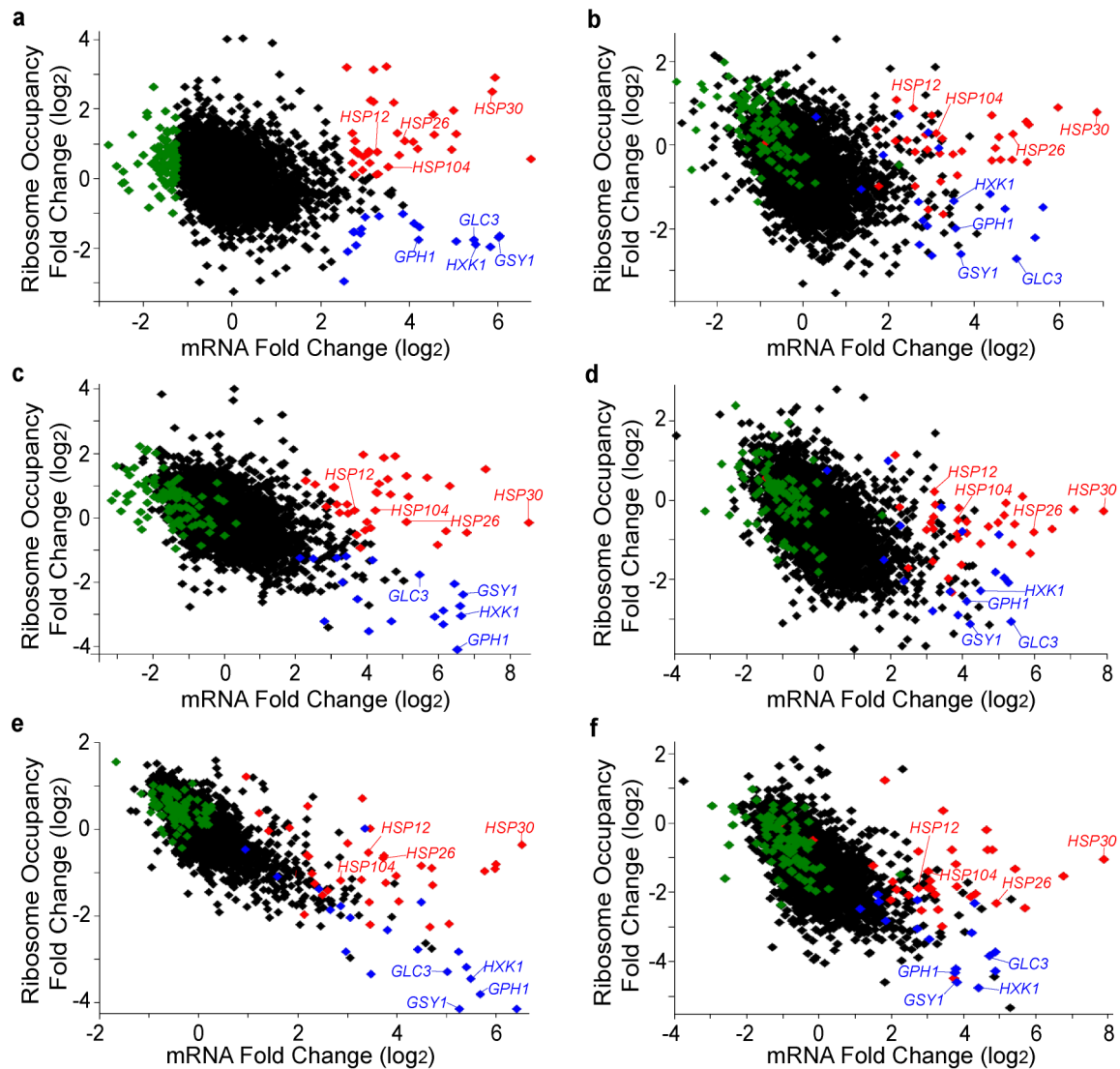
20. Haim-Vilmovsky, L. & Gerst, J. E. m-TAG: a PCR-based genomic integration method to visualize the localization of specific endogenous mRNAs *in vivo* in yeast. *Nature Protocols* **4**, 1274–1284 (2009).
21. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**, 343–345 (2009).
22. Longtine, M. S. *et al.* Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**, 953–961 (1998).
23. Hansen, A. S. & O'Shea, E. K. Promoter decoding of transcription factor dynamics involves a trade-off between noise and control of gene expression. *Mol. Syst. Biol.* **9**, 704 (2013).
24. Brandman, O. *et al.* A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. *Cell* **151**, 1042–1054 (2012).
25. Lutfiyya, L. L. *et al.* Characterization of three related glucose repressors and genes they regulate in *Saccharomyces cerevisiae*. *Genetics* **150**, 1377–1391 (1998).
26. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
27. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **9**, 62–66 (1979).
28. Ruxton, G. D. The unequal variance *t*-test is an underused alternative to Student's *t*-test and the Mann–Whitney *U* test. *Behav. Ecol.* **17**, 688–690 (2006).
29. Cook, K. E. & O'Shea, E. K. Hog1 controls global reallocation of RNA Pol II upon osmotic shock in *Saccharomyces cerevisiae*. *G3 (Bethesda)* **2**, 1129–1136 (2012).





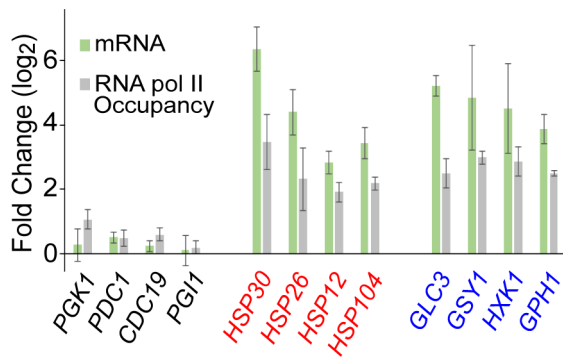
**Extended Data Figure 1 | Glucose starvation causes a reduction in overall translation, along with gene-specific changes in ribosome and mRNA read density.** **a**, Sedimentation profile of logarithmic-phase cells of strain BY4741 grown in SCD medium. The arrow marks the sedimentation of the 80S ribosome. **b**, Sedimentation profile of cells of strain BY4741 grown in SCD medium and then transferred to the same medium lacking glucose for 15 min. **c**, Ribosome and mRNA read density across the *HXK1* mRNA during logarithmic-phase growth in glucose-rich conditions and after 15 min of

glucose starvation. For the tracks labelled “mRNA”, the number of mRNA reads is shown normalized to the total number of sequence reads for that sample (in reads per million reads (RPM)). For the tracks labelled “Ribosome”, the number of ribosome reads is shown normalized to the total number of reads for that sample (in RPM). The initiation region was defined as a 36-base-pair (bp) region that contains 16 bp upstream of the AUG and 20 bp downstream. The downstream region is defined as the rest of the ORF.



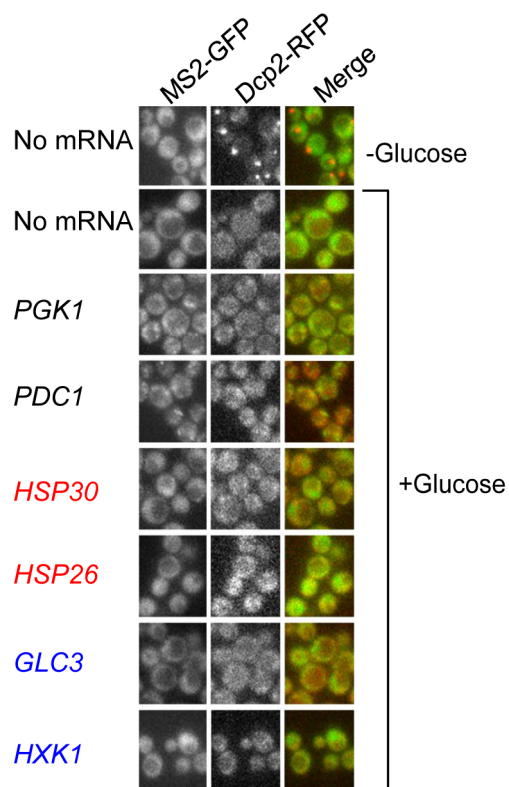
**Extended Data Figure 2 | Differences in ribosome occupancy of transcriptionally upregulated mRNAs upon glucose starvation are reproducible and independent of the mRNA isolation method.** Ribosomal profiling was performed on strains BY4741 and EY0690 grown in glucose-rich and glucose-starvation conditions. The fold change in ribosome occupancy versus the fold change in mRNA level 15 min after cells are transferred to medium lacking glucose is shown. Genes are represented by individual symbols on the plot. Ribosome occupancy was calculated for the coding region of each gene by dividing the total number of ribosome sequence counts in an ORF (normalized to the total number of aligned reads in RPM) by the number of mRNA sequence counts (RPM) in the same sequence. The coloured symbols in each panel show the gene classes defined from BY4741 ribosomal profiling of non-poly(A)<sup>+</sup>-selected mRNA replicate 1 in **a** (identical to Fig. 1a). Red symbols denote genes that have upregulated mRNA levels (>2.5) and higher ribosome occupancy (>0.09). Blue symbols denote genes that have upregulated mRNA levels (>2.5) with lower ribosome occupancy (<-1.0). Green symbols denote genes that have decreased mRNA levels (<-1.25) during glucose

limitation. While downregulated mRNAs are present at decreased levels, many of them have increased ribosome occupancy, and this subset of mRNAs is enriched for genes involved in ribosome biogenesis (26 of 84; false discovery rate (FDR)-adjusted  $P = 9.9 \times 10^{-4}$ ) by gene ontology analysis. Similarly, it has previously been observed that ribosome biogenesis mRNAs are present at decreased levels and have increased polysome association during early glucose starvation<sup>3</sup>. Black symbols represent all other genes in the genome for which measurements were obtained. The upregulated, higher ribosome occupancy genes (*HSP30*, *HSP26*, *HSP12* and *HSP104*) and the upregulated, lower ribosome occupancy genes (*GLC3*, *GSY1*, *GPH1* and *HXK1*) are labelled in each panel. **a**, BY4741 non-poly(A)<sup>+</sup>-selected RNA, ribosome profiling replicate 1 (same as Fig. 1a). **b**, EY0690 non-poly(A)<sup>+</sup>-selected RNA, ribosome profiling replicate 1. **c**, BY4741 poly(A)<sup>+</sup>-selected RNA, ribosomal profiling replicate 1. **d**, EY0690 poly(A)<sup>+</sup>-selected RNA, ribosomal profiling replicate 1. **e**, BY4741 poly(A)<sup>+</sup>-selected RNA, ribosomal profiling replicate 2. **f**, EY0690 poly(A)<sup>+</sup>-selected RNA, ribosomal profiling replicate 2.



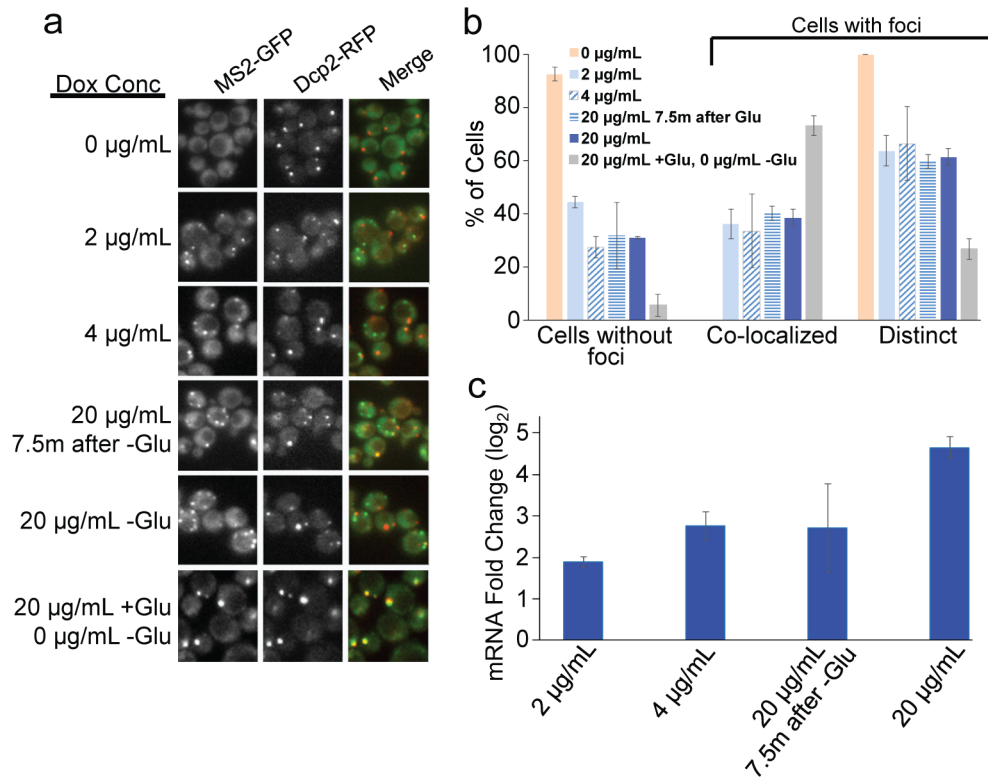
**Extended Data Figure 3 | The genes corresponding to upregulated mRNAs have increased RNA polymerase II occupancy upon glucose starvation.**

The mRNA levels of the indicated genes were measured by RNA sequencing (RNA-seq) after 15 min of glucose starvation, and these levels were divided by the levels obtained during growth in glucose-rich medium to obtain the fold-change values. The measurements were made on independent biological samples (with strains BY4741 and EY0690), and the values are presented as the mean  $\pm$  s.e.m. The RNA polymerase II occupancy was measured after 15 min of glucose starvation, and then this occupancy was divided by the occupancy in glucose-rich medium to obtain the fold-change values. RNA polymerase II occupancy was calculated from three independent biological replicates of BY4741, and the values are presented as the mean  $\pm$  s.e.m.



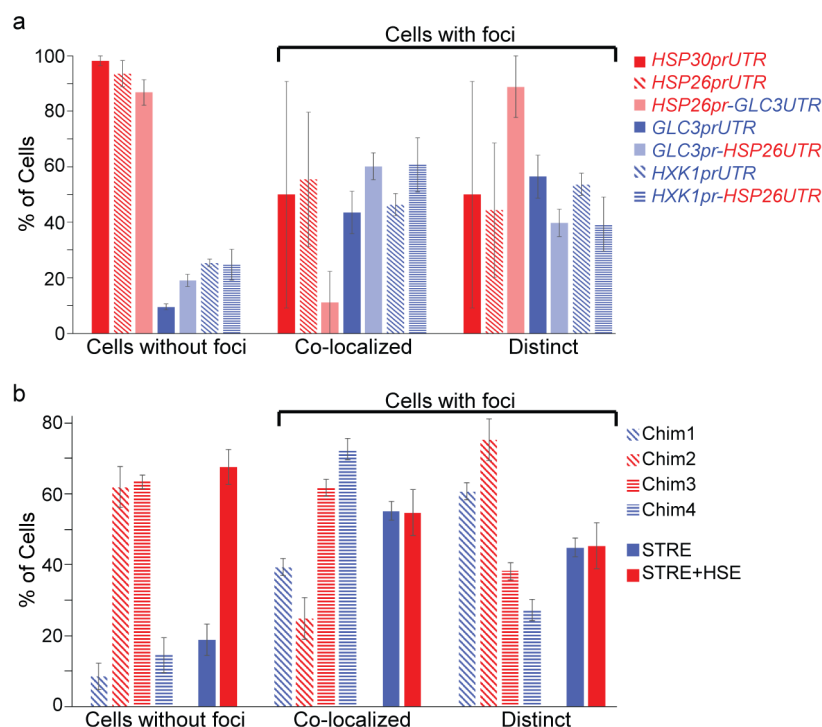
**Extended Data Figure 4 | Formation of mRNA foci is dependent on MS2-binding sites and does not occur in cells growing exponentially in medium containing glucose.** In the absence of mRNA containing MS2-binding sites, MS2-GFP remained diffusely localized during glucose starvation (first two rows, first column). When glucose was present in the medium, MS2 mRNA and the P body marker Dcp2-RFP were diffusely localized during the logarithmic phase of growth.





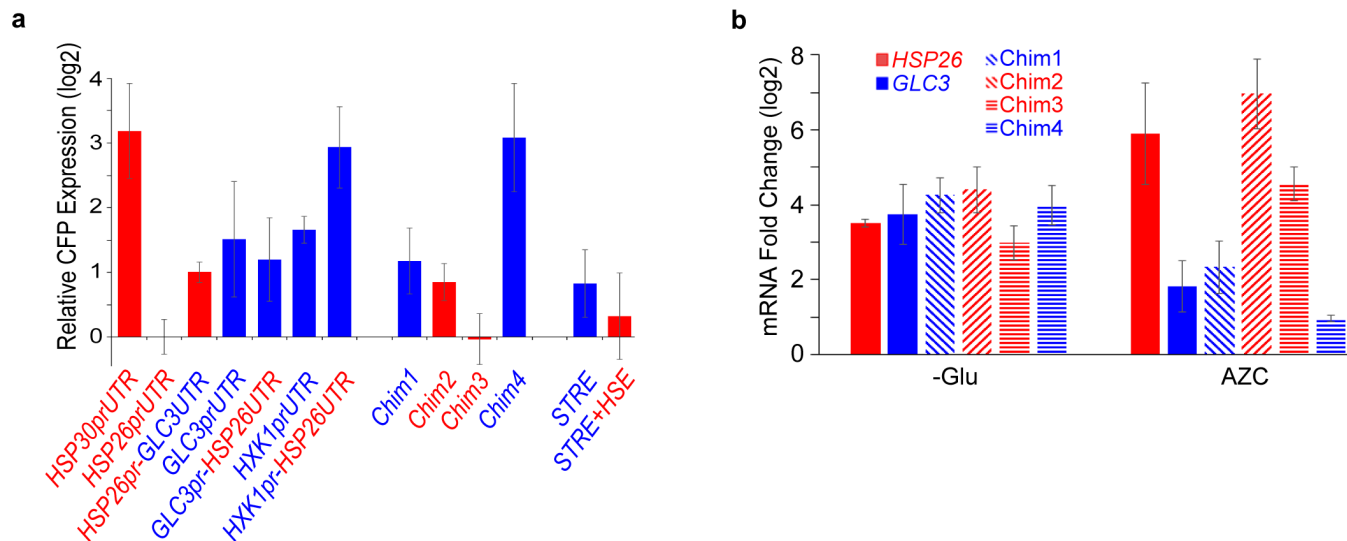
**Extended Data Figure 5 | Timing of *lacZ-MS2* induction relative to glucose starvation affects mRNA localization, whereas timing or level of induction during glucose starvation has no effect.** The expression of *lacZ-MS2* was either uninduced ( $0 \mu\text{g mL}^{-1}$ ), induced to different extents with varying concentrations of doxycycline during glucose starvation ( $2$ ,  $4$  or  $20 \mu\text{g mL}^{-1}$ ), induced at different times during glucose starvation ( $7.5$  min of glucose starvation with no doxycycline, then  $20 \mu\text{g mL}^{-1}$  doxycycline for the final  $7.5$  min of glucose starvation) or induced before glucose starvation ( $20 \mu\text{g mL}^{-1}$  doxycycline during logarithmic phase and none during glucose starvation) in the EY2897 strain. **a**, Localization of the mRNA was visualized using MS2-GFP after  $15$  min of glucose starvation for all strains. Dcp2-RFP was used to visualize P body localization. *lacZ-MS2* expression before glucose starvation

caused high co-localization with P bodies, while mRNA induction during glucose starvation caused the formation of mRNA foci that both co-localized with P bodies and were distinct from P bodies. **b**, Quantification of the localization data shown in **a**. The values are presented as the mean  $\pm$  s.e.m. and were calculated as follows: in quadruplicate (two biological replicates with two technical replicates per sample) for  $0$ ,  $2$ ,  $20 \mu\text{g mL}^{-1}$  doxycycline in the absence of glucose and  $20 \mu\text{g mL}^{-1}$  doxycycline in the presence of glucose; and in triplicate on technical replicates for  $4$  and  $20 \mu\text{g mL}^{-1}$  doxycycline  $7.5$  min after the removal of glucose. **c**, Quantification of *lacZ-MS2* mRNA levels  $15$  min after glucose starvation. The fold change was calculated compared with the uninduced sample ( $0 \mu\text{g mL}^{-1}$  doxycycline) and normalized to *ACT1* abundance, as calculated from three independent biological replicates.



**Extended Data Figure 6 | Promoter sequences determine mRNA localization upon glucose starvation.** **a**, The promoter and 5' UTR of the indicated genes were fused upstream of *CFP-MS2* in the plasmid pRS305 and integrated into EY0690. The mRNA localization was measured after 15 min of glucose starvation. The values are presented as the mean  $\pm$  s.e.m. from Fig. 3c as calculated on a minimum of 30 cells in quadruplicate (two biological

replicates with two technical replicates per sample). **b**, Localization of *CFP-MS2* mRNAs driven by chimaeric *HSP26* and *GLC3* promoters or synthetic (STRE or STRE+HSE) promoters upon glucose starvation. The values are presented as the mean  $\pm$  s.e.m. from Fig. 4b as calculated on a minimum of 25 cells in quadruplicate (two biological replicates with two technical replicates per sample).



**Extended Data Figure 7 | mRNA levels of *CFP-MS2* under different conditions controlled by various promoter-UTR combinations.** **a**, The relative levels of *CFP-MS2* mRNA, under the control of the indicated promoter and UTR regions 15 min after glucose starvation, as measured by qPCR. The values are normalized to *ACT1* abundance and are presented as the mean  $\pm$  s.e.m. relative to the *HSP26prUTR-CFP* levels and calculated from

three independent biological replicates. **b**, Fold change in *CFP-MS2* mRNA abundance after 15 min of glucose starvation ( $-Glu$ ) or after treatment with 10 mM AZC for 2 h ( $+AZC$ ), relative to levels during logarithmic-phase growth in glucose-rich medium. *CFP-MS2* mRNA was measured by qPCR and normalized to *ACT1* levels. The values are presented as the mean  $\pm$  s.e.m. as calculated from three independent biological replicates.

Mig1-binding element  
 4xSTRE element  
 3xHSE element  
 Attenuated *cyc1* promoter  
 ATG of CFP

#### 4xSTRE

TAAAAATGCGGGGATTATATCAGTTATTACCCTCGAGAATTGGTAAGGG  
 CCCAATTGGTAAGGGGCCAATTGGTAAGGGGCCAATTGGTAAGGGGCC  
 TCGAGCAGATCCGCCAGGCGTGATATAGCGTGGATGGCCAGGCAACT  
 TTAGTGCTGACACATACAGGCATATATATGTGTGCGACGACACATG  
 ATCATATGGCATGCATGTGCTCTGTATGTATATAAACTCTTGTTTTCTT  
 CTTTTCTCTAAATATTCTTTCCTTATACATTAGGTCCTTTGTAGCATAAA  
 TTAATACTTCTATAGACACGCAAACACAAATACACACACTAAATTAA  
 TAATG

#### 4xSTRE+3xHSE

TAAAAATGCGGGGATTATATCAGTTATTACCCTCGAGAATTGGTAAGGG  
 CCCAATTGGTAAGGGGCCAATTGGTAAGGGGCCAATTGGTAAGGGGCC  
 TCGAGCAGACCCTCGAGCTAGAAGCTTCTAGAAGCTTCTAGAGGATCC  
 CCGTCGAGCAGATCCGCCAGGCGTGATATAGCGTGGATGGCCAGGCA  
 ACTTTAGTGCTGACACATACAGGCATATATATGTGTGCGACGACACA  
 TGATCATATGGCATGCATGTGCTCTGTATGTATATAAACTCTTGTTTTCT  
 TTTCTTCTCTAAATATTCTTTCCTTATACATTAGGTCCTTTGTAGCATA  
 AATACTTCTATAGACACGCAAACACAAATACACACACTAAATT  
 AATAATG

**Extended Data Figure 8 | Synthetic STRE±HSE promoter sequences.** The STRE±HSE elements were placed upstream of an attenuated *CYC1* promoter<sup>17</sup> driving the expression of *CFP-MS2*. A Mig1-binding element was included

upstream of the promoter elements to reduce expression pre-starvation. The Mig1-binding element is shown in grey; the 4× STRE is labelled in blue; the 3× HSE is labelled in red; and the *CYC1* promoter is labelled in yellow.



Extended Data Table 1 | Gene ontology analysis of the classes of genes that are differentially regulated in glucose starvation

BY4741			EY0690		
Upregulated Higher-Ribo n=26	Genes	p-value	Upregulated Higher-Ribo n=36	Genes	p-value
Response to temperature stimulus	14	1.3E-9	Response to abiotic stimulus	21	9.1E-12
Response to abiotic stimulus	16	2.4E-9	Response to temperature stimulus	18	1.7E-11
Cellular response to heat	11	1.4E-6	Cellular response to heat	15	9.0E-9
Vacuolar protein catabolic process	8	1.7E-3	Cellular response to stress	16	3.4E-3
Upregulated Lower-Ribo n=18			Upregulated Lower-Ribo n=37		
Glucose metabolic process	7	7.8E-4	Vacuolar protein catabolic process	10	1.4E-4
Vacuolar protein catabolic process	7	2.0E-3	Energy reserve metabolic process	7	5.9E-4
Hexose metabolic process	7	2.5E-3	Glycogen metabolic process	6	2.6E-3
			Glucose metabolic process	8	8.2E-3
Downregulated n=84			Downregulated n=83		
RNA modification	19	7.6E-10	Ribosome biogenesis	33	1.6E-13
ncRNA metabolic process	25	2.3E-5	Ribonucleoprotein complex biogenesis	33	8.4E-12
rRNA processing	19	9.0E-5	rRNA processing	24	3.9E-9
Ribonucleoprotein complex biogenesis	24	1.3E-4	maturation of SSU-rRNA	16	6.7E-9
Ribosome biogenesis	26	9.9E-4	ncRNA metabolic process	27	5.9E-7
RNA processing	8	1.1E-3	RNA processing	28	4.3E-5
Methionine biosynthetic process	8	1.1E-3	RNA modification	13	7.3E-4
Sulfur amino acid biosynthetic process	8	3.2E-3	maturation of 5.8S rRNA	10	2.5E-3

DAVID analysis software was used to find the Gene Ontology (GO) terms that were significantly enriched (FDR-adjusted  $P$  value  $< 1.0 \times 10^{-2}$ ) in differentially regulated groups of genes from non-poly(A)<sup>+</sup>-selected RNA-seq data and ribosomal profiling replicate 1 data for each strain (mRNA upregulated, higher ribosome occupancy (red); mRNA upregulated, lower ribosome occupancy (blue); mRNA downregulated (green)). GO terms that were common to BY4741 and EY0690 are shown in bold.

Extended Data Table 2 | Transcription start sites of mRNAs produced from promoters driving differential localization and protein production

<i>HSP26prUTR</i>	AAAGCAAACAAACAACTAAACAAATTAACATG ATTAAACAGGTATCCAAAAAGCAAACAAACAACTAAACAAATTAACATG ATTAAACAGGTATCCAAAAAGCAAACAAACAACTAAACAAATTAACATG ATTAAACAGGTATCCAAAAAGCAAACAAACAACTAAACAAATTAACATG ATATCAGATCTCTATTAACAGGTATCCAAAAAGCAAACAAACAACTAAACAAATTAACATG
<i>GLC3pr-HSP26UTR</i>	TAAACAGGTATCCAAAAAGCAAACAAACAACTAAACAAATTAACATG ATTAAACAGGTATCCAAAAAGCAAACAAACAACTAAACAAATTAACATG ATTAAACAGGTATCCAAAAAGCAAACAAACAACTAAACAAATTAACATG GATCTCTATTAACAGGTATCCAAAAAGCAAACAAACAACTAAACAAATTAACATG
<i>HXK1pr-HSP26UTR</i>	ATTAAACAGGTATCCAAAAAGCAAACAAACAACTAAACAAATTAACATG ATTAAACAGGTATCCAAAAAGCAAACAAACAACTAAACAAATTAACATG ATTAAACAGGTATCCAAAAAGCAAACAAACAACTAAACAAATTAACATG TATCAGATCTCTATTAACAGGTATCCAAAAAGCAAACAAACAACTAAACAAATTAACATG
<i>GLC3pr-GLC3UTR</i>	AAGTATAAGAACCGTCAAGAATAAAATG AAGTATAAGAACCGTCAAGAATAAAATG AAGTATAAGAACCGTCAAGAATAAAATG AAACCAAGTATAAGAACCGTCAAGAATAAAATG AAACCAAGTATAAGAACCGTCAAGAATAAAATG
<i>HSP26UTR-GLC3UTR</i>	CAAACCAAGTATAAGAACCGTCAAGAATAAAATG ACAAACCAAGTATAAGAACCGTCAAGAATAAAATG GATAAA CAAACCAAGTATAAGAACCGTCAAGAATAAAATG ACCGATAAA CAAACCAAGTATAAGAACCGTCAAGAATAAAATG ACCGATAAA CAAACCAAGTATAAGAACCGTCAAGAATAAAATG
<i>STRE</i>	ACGCAAACACAAATACACACTAAATTAATAATG ATAGACACGCAAACACAAATACACACTAAATTAATAATG ATAGACACGCAAACACAAATACACACTAAATTAATAATG ATACTTCTATAGACACGCAAACACAAATACACACTAAATTAATAATG ATACTTCTATAGACACGCAAACACAAATACACACTAAATTAATAATG GTAGCATAAATTACTTACTATAGACACGCAAACACAAATACACACTAAATTAATAATG
<i>STRE+HSE</i>	AAATTAATAATG AAACACAAATACACACTAAATTAATAATG ATACTTCTATAGACACGCAAACACAAATACACACTAAATTAATAATG ATACTTCTATAGACACGCAAACACAAATACACACTAAATTAATAATG ATACTTATATAGACACGCAAACACAAATACACACTAAATTAATAATG ATAAATTACTTACTTCTATAGACACGCAAACACAGATACACACTAAATTAATAATG

5' Rapid amplification of cDNA ends (RACE) was used to determine the transcriptional start sites of the *CFP-MS2* mRNAs driven by the indicated promoter-5' UTR combinations.

# Sae2 promotes dsDNA endonuclease activity within Mre11–Rad50–Xrs2 to resect DNA breaks

Elda Cannavo<sup>1</sup> & Petr Cejka<sup>1</sup>

To repair double-strand DNA breaks by homologous recombination, the 5'-terminated DNA strand must first be resected, which generates 3' single-stranded DNA overhangs. Genetic evidence suggests that this process is initiated by the Mre11–Rad50–Xrs2 (MRX) complex<sup>1–3</sup>. However, its involvement was puzzling, as the complex possesses exonuclease activity with the opposite (3' to 5') polarity from that required for homologous recombination<sup>4,5</sup>. Consequently, a bidirectional model has been proposed<sup>6–8</sup> whereby dsDNA is first incised endonucleolytically and MRX then proceeds back to the dsDNA end using its 3' to 5' exonuclease. The endonuclease creates entry sites for Sgs1–Dna2 and/or Exo1, which then carry out long-range resection in the 5' to 3' direction. However, the identity of the endonuclease remained unclear. Using purified *Saccharomyces cerevisiae* proteins, we show that Sae2 promotes dsDNA-specific endonuclease activity by the Mre11 subunit within the MRX complex. The endonuclease preferentially cleaves the 5'-terminated dsDNA strand, which explains the polarity paradox. The dsDNA end clipping is strongly stimulated by protein blocks at the DNA end, and requires the ATPase activity of Rad50 and physical interactions between MRX and Sae2. Our results suggest that MRX initiates dsDNA break processing by dsDNA endonuclease rather than exonuclease activity, and that Sae2 is the key regulator of this process. These findings demonstrate a probable mechanism for the initiation of dsDNA break processing in both vegetative and meiotic cells.

Recombinant MRX is a 3'–5' exonuclease that requires manganese and not magnesium (Extended Data Fig. 1a–d), as shown previously<sup>4</sup>. Sae2 was prepared by adding a sequence coding for the maltose-binding protein (MBP) to its amino terminus to allow us to purify the protein to near homogeneity; the MBP tag was cleaved off during later purification steps (Extended Data Fig. 1e, f). To study the interplay of the recombinant MRX and Sae2 proteins, we used a dsDNA substrate with a biotin-streptavidin block at one of its DNA ends (Fig. 1a). As expected, MRX alone released the <sup>32</sup>P label from the 3'-terminated strand on the unprotected side of the substrate (Fig. 1a, lanes 3–6). Recombinant Sae2 had no activity (Fig. 1a, lanes 11–14). To our surprise, when both MRX and Sae2 were present, a novel degradation band appeared, indicating nucleolytic cleavage in approximately the centre of the dsDNA (Fig. 1a, lanes 7–10, Extended Data Fig. 2a). This suggested that, when combined, MRX and Sae2 proteins acquire the capacity to cleave dsDNA endonucleolytically. The amount of endonucleolytic product increased with Sae2 concentration and was not affected by the single-stranded DNA binding protein, RPA (Extended Data Fig. 2b–d). To establish whether the apparent endonuclease is inherent to MRX or Sae2, we purified a variant of MRX<sup>9</sup> containing nuclease-deficient Mre11(H125L;D126V), designated MRX-nd (Extended Data Fig. 3a, b). The mutant was deficient in the Sae2-promoted endonuclease activity; thus, both exo- and endonuclease products are dependent on the same Mre11 active site (Fig. 1b). Under our conditions, Sae2 exhibited no nuclease activity (Extended Data Fig. 4a–e)<sup>10</sup>.

ATP binding and hydrolysis bring about conformational changes within the Rad50 subunit of MRX, which has emerged as a critical regulator of MRX-dependent functions<sup>11,12</sup>. We found that ATP was essential

for the MRX- and Sae2-dependent clipping of dsDNA (Fig. 1c). Neither non-hydrolysable ATP analogue ATP $\gamma$ S nor ADP supported the endonuclease (Extended Data Fig. 5a). MRX variants deficient in ATP binding and/or hydrolysis due to mutations in the Walker A-type motif of the Rad50 subunit<sup>13</sup> (K40A and K40R) were endonuclease-deficient even in the presence of ATP (Fig. 1d, Extended Data Fig. 3c). Previously, both exo- and endonuclease activities of MRX were found to require manganese<sup>4</sup>, and the crystal structure of Mre11 was shown to contain two manganese ions<sup>14</sup>; little, if any, activity was observed in the presence of more physiological magnesium. To our surprise, we found that magnesium-only conditions supported, although only to a limited extent, Sae2-dependent dsDNA clipping (Fig. 1e, lane 6). Almost no exonuclease products were detected. Magnesium and manganese together maximally supported both exo- and endonuclease activities (Fig. 1e, lane 10). This is in agreement with magnesium being required for the ATPase of Rad50 and manganese optimally promoting the nuclease of Mre11. We also show that an excess of magnesium relative to manganese was required for the stimulation of the MRX endonuclease (Extended Data Fig. 5b). In meiosis, homologous recombination is initiated by the formation of double strand breaks catalysed by Spo11, which remains covalently attached to the DNA ends<sup>15</sup>. Yeast *rad50s* mutants fail to produce Spo11–oligonucleotide complexes, suggesting that Rad50 is critical for the anticipated endonucleolytic initiation of dsDNA break processing<sup>16,17</sup>. We show here that the Rad50S MRX variant, carrying the single K81I amino acid substitution, completely lacks the capacity to clip dsDNA (Fig. 1f), which is likely to provide a mechanistic explanation of the *rad50s* mutant phenotypes. *rad50s* mutants often closely resemble *sae2Δ* cells<sup>18</sup>, which agrees with our observation that Sae2 promotes MRX endonuclease activity.

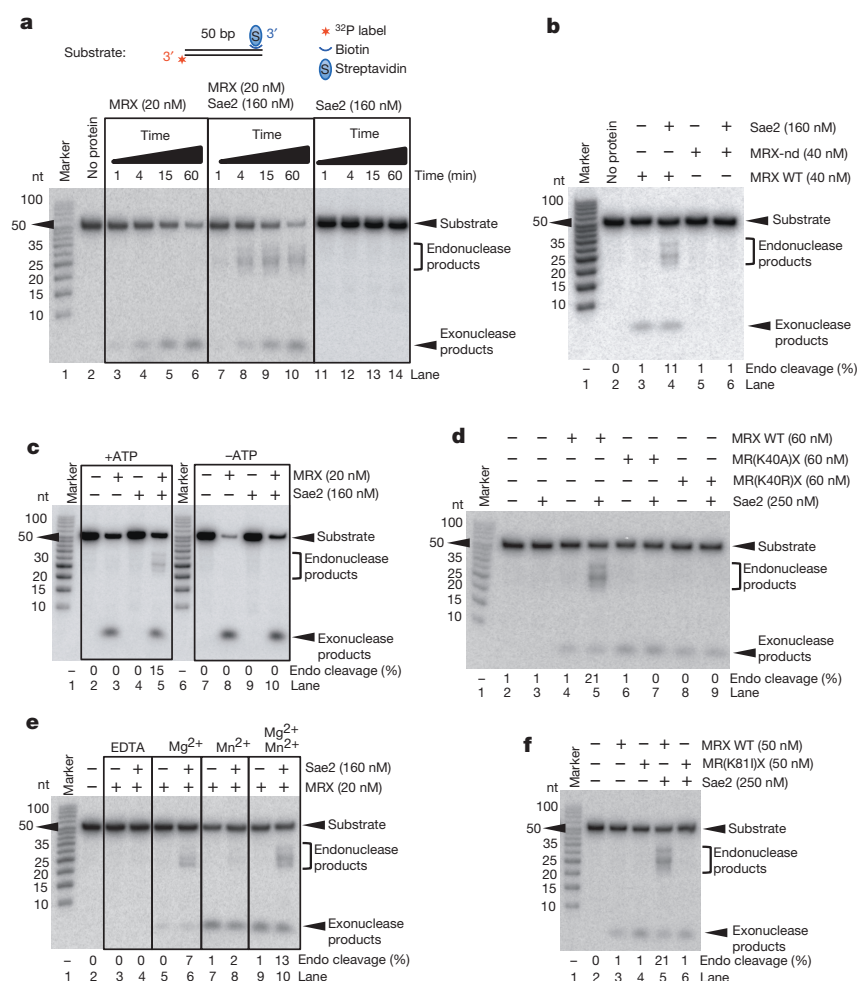
The requirement for MRX and Sae2 in promoting resection on DNA ends bound by Ku, aborted topoisomerases or Spo11 suggests that the Mre11 nuclease has a more general role in the processing of blocked or modified DNA ends<sup>17,19–23</sup>, which may not require specific interactions between MRX and the protein block. We next used a DNA substrate biotinylated on both DNA strands near both 5' and 3' termini. In the absence of streptavidin, MRX removed the 3' label via its 3'–5' exonuclease independently of Sae2, and almost no endonuclease fragments were detected (Fig. 2, lanes 8 and 10; see also Extended Data Fig. 6a). In the presence of streptavidin, the exonucleolytic degradation of this fully blocked substrate was completely inhibited. Instead, we observed robust endonucleolytic cleavage when MRX and Sae2 were combined (Fig. 2, lane 5; see also Extended Data Fig. 6b–f). Avidin acted similarly to streptavidin; furthermore, protein blocks did not promote endonucleolytic cleavage when Sae2 was combined with Mre11, Mre11–Xrs2, Exo1 or Dna2 nucleases (Extended Data Fig. 6g–i). Structure-specific MRX endonuclease activity can be observed on circular ssDNA, but Sae2 did not stimulate MRX cleavage of this substrate (Extended Data Fig. 7), showing that Sae2 specifically promotes the MRX endonuclease in the vicinity of blocked DNA ends.

Homologous recombination is initiated by 5' DNA end resection that leaves the prerequisite 3' ssDNA tails. Genetic experiments indicated that MRX is part of a complex that resects DNA with this polarity<sup>2</sup>; however, these results were incompatible with the observed 3'–5'

<sup>1</sup>Institute of Molecular Cancer Research, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.

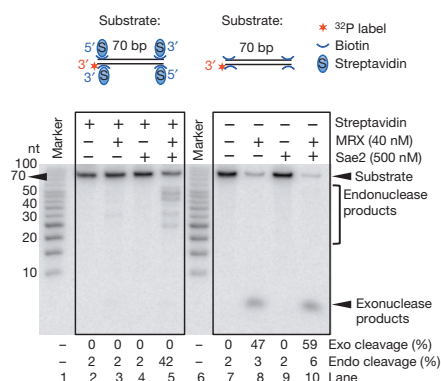
**Figure 1 | Sae2 promotes dsDNA endonuclease activity within the Mre11 subunit of MRX.**

**a**, Concerted action of Sae2 and MRX results in a vigorous dsDNA endonuclease activity. nt, nucleotides. **b**, The endonuclease activity is dependent on the Mre11 nuclease. Wild-type or nuclease-deficient MRX and/or Sae2 were used, as indicated. Endo cleavage (%), average percentage of endonucleolytic products, on the basis of two independent experiments. **c**, Endonucleolytic activity of MRX–Sae2 is dependent on ATP. **d**, Sae2 does not activate endonuclease of MRX mutants deficient in ATP hydrolysis. **e**, Nuclease activity of MRX–Sae2 and its dependence on magnesium (5 mM) and manganese (5 mM). **f**, Sae2 does not activate endonuclease activity of the Rad50S variant of MRX.

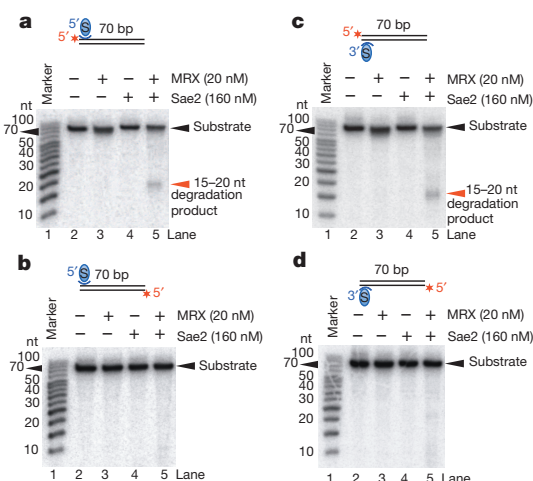


exonuclease activity of the recombinant factor<sup>4</sup>. We wondered whether the endonuclease capacity described here could explain the apparent polarity paradox. To determine this, we used 70-base-pair long dsDNA substrates that contained a single biotin–streptavidin block either at the 5' or 3'-terminated strands, and placed a  $^{32}$ P label at various positions to monitor the sites of endonucleolytic cleavage. MRX and Sae2 cleaved DNA approximately 15–20 nucleotides away from the protein-blocked dsDNA end (Fig. 3a, c). MRX activity was not differentially

affected by the attachment of streptavidin to either the 3' or 5' terminated strand; this is probably owing to the larger size of the protein with respect to the diameter of dsDNA. The cleavage of the 5'-terminated strand, however, was clearly preferred (Fig. 3a–d). This was further confirmed using a 100-bp-long oligonucleotide as well as plasmid-length



**Figure 2 | Sae2-dependent MRX endonuclease activity is stimulated by a protein block at the dsDNA end.** Nuclease assay was performed with biotin-labelled dsDNA, with or without streptavidin, as indicated. Protein block inhibited the exonuclease activity but promoted the endonuclease activity of MRX–Sae2. Endo or exo cleavage (%), average percentage of endo or exo products, on the basis of two independent experiments.



**Figure 3 | Sae2-dependent MRX endonuclease activity is specific to the 5'-terminated DNA strand.** **a–d**, Nuclease assay was performed with streptavidin-blocked dsDNA at either the 5' or the 3' end, with  $^{32}$ P-label at various positions, as indicated. MRX–Sae2 preferentially incised the 5'-terminated strand about 15–20 nucleotides from the DNA end.



DNA substrate (Extended Data Fig. 8)<sup>24</sup>. Together, these data indicate that MRX is likely to bind dsDNA directionally, and proper orientation of the MRX complex with respect to the protein block results in the observed polarity of DNA cleavage, which is prerequisite for homologous recombination.

Previous attempts to demonstrate an interaction between MRX and Sae2 *in vivo* or *in vitro* under native conditions were unsuccessful<sup>10,25</sup>. The functional interaction between MRX and Sae2 proteins observed here prompted us to revisit this issue using MBP-tagged Sae2 protein immobilized on amylose resin and recombinant MRX complex. As shown in Fig. 4a, lane 4, the Sae2 eluate contained all subunits of the MRX complex. In contrast, MRX did not bind to amylose-bound MBP tag, showing that MRX does not interact with either amylose resin or the MBP affinity tag non-specifically (Fig. 4a, lane 5). We attempted to test which component of the MRX complex mediates the interaction, but we could only detect very weak binding of Sae2 to both Mre11 and Xrs2 subunits (Extended Data Fig. 9), and we failed to detect interaction with Rad50 (not shown). Thus, Sae2 is likely to bind several subunits of the MRX complex.

We prepared a variety of Sae2 mutants on the basis of conserved residues, mutant phenotypes and phosphorylation sites<sup>10,18,26–29</sup> (Extended Data Fig. 10a–e). Sae2 mutant lacking the first 169 amino acids (170–345) could still promote MRX endonuclease activity, whereas a mutant lacking the last 95 amino acids (1–250) was inactive, showing that the carboxy-terminal part of Sae2 is critical for the stimulation of the MRX endonuclease. In accord, point mutations at the N terminus did not affect Sae2 function. One of these, Sae2(E24V), was shown to exhibit a severe

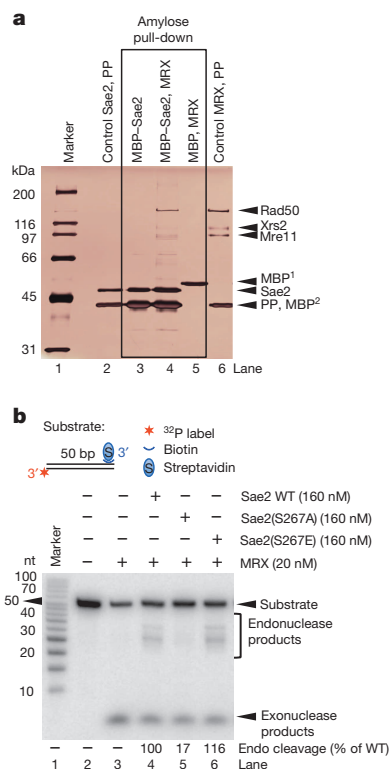
hairpin-processing defect, but behaved as wild type with regard to meiotic progression<sup>26</sup>. In agreement, we found no synergy in hairpin cleavage when both MRX and Sae2 were combined (Extended Data Fig. 4d, e), showing that the dsDNA clipping mechanism described here does not explain hairpin cleavage, whereas it is probably relevant for understanding the processing of protein-blocked DNA ends. In contrast, conserved residues in the C-terminal part of Sae2 were important for its function. In particular, mutations of residues in the region between amino acids 267 and 279 into alanines rendered Sae2 almost completely inactive, and mutations at positions 264 and 300 resulted in severe inhibition (Extended Data Fig. 10c–e). The ability to stimulate MRX endonuclease activity did not correlate with the capacity of the Sae2 variants to bind DNA; instead, we found that Sae2 mutants that failed to stimulate MRX were often impaired in their interaction with the heterotrimer (Extended Data Fig. 10f, g). This suggests that the C-terminal region of Sae2 between residues 267 and 280 either directly interacts with MRX or is important for proper folding of Sae2. Phosphorylation of the conserved Ser 267 residue of Sae2 by CDK is required for all of its functions *in vivo*<sup>27</sup>. We found that non-phosphorylatable Sae2(S267A) mutant was deficient in its capacity to promote MRX endonuclease, whereas phospho-mimicking Sae2(S267E) mutant was indistinguishable from wild type (Fig. 4b). This confirms the critical importance of Ser 267 in Sae2. Treatment of our recombinant Sae2 with protein phosphatase 1 strongly reduced its ability to promote MRX endonuclease (Extended Data Fig. 10h), suggesting that Sae2 protein is being phosphorylated in Sf9 cells, and this modification is important for its capacity to activate MRX. This might also explain why the stimulatory function of Sae2 went undetected in a previous study that used *Escherichia coli* expression<sup>10</sup>. Finally, we note that none of the Sae2 mutants tested affected the exonuclease activity of MRX (Fig. 4b, Extended Data Fig. 10a–e).

Our results offer a mechanistic explanation of how double strand break processing can be initiated by the MRX complex. We have shown that Sae2 promotes dsDNA-specific endonuclease within the Mre11 subunit of MRX in the vicinity of protein blocks at a dsDNA end. The data suggest that direct and species-specific interaction between MRX–Sae2 and the protein block is not required; however, we believe that any such interaction might facilitate recruitment and further promote the Mre11 endonuclease. The endonuclease activity of MRX preferentially clips the 5′-terminated DNA strand. This polarity is required for homologous recombination and would generate 3′ tailed substrates that are optimal for Exo1 and/or Sgs1–Dna2. These data thus directly support the bidirectional resection model<sup>6–8</sup>.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 27 May; accepted 15 August 2014.

Published online 17 September 2014.



**Figure 4 | Sae2 physically and functionally interacts with MRX.** **a**, Sae2 interacts with MRX. MBP-tagged Sae2 was bound to amylose resin and incubated (lane 4) or not (lane 3) with recombinant MRX. MRX does not bind to amylose-bound MBP tag alone (lane 5). Controls: recombinant Sae2 (125 ng, lane 2), MRX (250 ng, lane 6). All samples were treated with PreScission protease before gel analysis to distinguish MRX–Sae2 from Mre11. MBP<sup>1</sup>, maltose-binding protein expressed in *E. coli*; MBP<sup>2</sup>, maltose-binding protein resulting from cleavage with PreScission protease; PP, PreScission protease. **b**, Nuclease assay was performed with MRX and Sae2 variants, as indicated. Endo cleavage (% of WT), average percentage of endonucleolytic products, normalized to wild-type Sae2, on the basis of two independent experiments.

- Mimitou, E. P. & Symington, L. S. Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature* **455**, 770–774 (2008).
- Zhu, Z., Chung, W. H., Shim, E. Y., Lee, S. E. & Ira, G. Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. *Cell* **134**, 981–994 (2008).
- Gravel, S., Chapman, J. R., Magill, C. & Jackson, S. P. DNA helicases Sgs1 and BLM promote DNA double-strand break resection. *Genes Dev.* **22**, 2767–2772 (2008).
- Trujillo, K. M. & Sung, P. DNA structure-specific nuclease activities in the *Saccharomyces cerevisiae* Rad50–Mre11 complex. *J. Biol. Chem.* **276**, 35458–35464 (2001).
- Trujillo, K. M. *et al.* Yeast Xrs2 binds DNA and helps target Rad50 and Mre11 to DNA ends. *J. Biol. Chem.* **278**, 48957–48964 (2003).
- Shibata, A. *et al.* DNA double-strand break repair pathway choice is directed by distinct MRE11 nuclease activities. *Mol. Cell* **53**, 7–18 (2014).
- Garcia, V., Phelps, S. E., Gray, S. & Neale, M. J. Bidirectional resection of DNA double-strand breaks by Mre11 and Exo1. *Nature* **479**, 241–244 (2011).
- Zakharyevich, K. *et al.* Temporally and biochemically distinct activities of Exo1 during meiosis: double-strand break resection and resolution of double Holliday junctions. *Mol. Cell* **40**, 1001–1015 (2010).
- Bressan, D. A., Olivares, H. A., Nelms, B. E. & Petrini, J. H. Alteration of N-terminal phosphoesterase signature motifs inactivates *Saccharomyces cerevisiae* Mre11. *Genetics* **150**, 591–600 (1998).

10. Lengsfeld, B. M., Rattray, A. J., Bhaskara, V., Ghirlando, R. & Paull, T. T. Sae2 is an endonuclease that processes hairpin DNA cooperatively with the Mre11/Rad50/Xrs2 complex. *Mol. Cell* **28**, 638–651 (2007).
11. Deshpande, R. A. *et al.* ATP-driven Rad50 conformations regulate DNA tethering, end resection, and ATM checkpoint signaling. *EMBO J.* **33**, 482–500 (2014).
12. Lammens, K. *et al.* The Mre11:Rad50 structure shows an ATP-dependent molecular clamp in DNA double-strand break repair. *Cell* **145**, 54–66 (2011).
13. Chen, L. *et al.* Effect of amino acid substitutions in the Rad50 ATP binding domain on DNA double strand break repair in yeast. *J. Biol. Chem.* **280**, 2620–2627 (2005).
14. Hopfner, K. P. *et al.* Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. *Cell* **105**, 473–485 (2001).
15. Keeney, S., Giroux, C. N. & Kleckner, N. Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* **88**, 375–384 (1997).
16. Alani, E., Padmore, R. & Kleckner, N. Analysis of wild-type and *rad50* mutants of yeast suggests an intimate relationship between meiotic chromosome synapsis and recombination. *Cell* **61**, 419–436 (1990).
17. Neale, M. J., Pan, J. & Keeney, S. Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature* **436**, 1053–1057 (2005).
18. Rattray, A. J., McGill, C. B., Shafer, B. K. & Strathern, J. N. Fidelity of mitotic double-strand-break repair in *Saccharomyces cerevisiae*: a role for *SAE2/COM1*. *Genetics* **158**, 109–122 (2001).
19. Langerak, P., Mejia-Ramirez, E., Limbo, O. & Russell, P. Release of Ku and MRN from DNA ends by Mre11 nuclease activity and Ctp1 is required for homologous recombination repair of double-strand breaks. *PLoS Genet.* **7**, e1002271 (2011).
20. Mimitou, E. P. & Symington, L. S. Ku prevents Exo1 and Sgs1-dependent resection of DNA ends in the absence of a functional MRX complex or Sae2. *EMBO J.* **29**, 3358–3369 (2010).
21. Shim, E. Y. *et al.* *Saccharomyces cerevisiae* Mre11/Rad50/Xrs2 and Ku proteins regulate association of Exo1 and Dna2 with DNA breaks. *EMBO J.* **29**, 3370–3380 (2010).
22. Liu, C., Pouliot, J. J. & Nash, H. A. Repair of topoisomerase I covalent complexes in the absence of the tyrosyl-DNA phosphodiesterase Tdp1. *Proc. Natl Acad. Sci. USA* **99**, 14970–14975 (2002).
23. Hartsuiker, E., Neale, M. J. & Carr, A. M. Distinct requirements for the Rad32(Mre11) nuclease and Ctp1(CtIP) in the removal of covalently bound topoisomerase I and II from DNA. *Mol. Cell* **33**, 117–123 (2009).
24. Hopkins, B. B. & Paull, T. T. The *P. furiosus* Mre11/Rad50 complex promotes 5' strand resection at a DNA double-strand break. *Cell* **135**, 250–260 (2008).
25. Uetz, P. *et al.* A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature* **403**, 623–627 (2000).
26. Kim, H. S. *et al.* Functional interactions between Sae2 and the Mre11 complex. *Genetics* **178**, 711–723 (2008).
27. Huertas, P., Cortes-Ledesma, F., Sartori, A. A., Aguilera, A. & Jackson, S. P. CDK targets Sae2 to control DNA-end resection and homologous recombination. *Nature* **455**, 689–692 (2008).
28. Akamatsu, Y. *et al.* Molecular characterization of the role of the *Schizosaccharomyces pombe* *nip1<sup>+</sup>/ctp1<sup>+</sup>* gene in DNA double-strand break repair in association with the Mre11-Rad50-Nbs1 complex. *Mol. Cell. Biol.* **28**, 3639–3651 (2008).
29. Penkner, A. *et al.* A conserved function for a *Caenorhabditis elegans* Com1/Sae2/CtIP protein homolog in meiotic recombination. *EMBO J.* **26**, 5071–5082 (2007).

**Acknowledgements** We thank L. Ranjha, R. Anand, L. Mlejnkova and M. Levikova for assistance and comments on the manuscript. This work was supported by Swiss National Science Foundation Grant PP00P3 133636 to P.C.

**Author Contributions** P.C. and E.C. designed the study. E.C. and P.C. carried out the experiments. P.C. and E.C. wrote the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.C. ([cejka@imcr.uzh.ch](mailto:cejka@imcr.uzh.ch)).

## METHODS

**Recombinant proteins.** The *SAE2* gene was amplified from genomic *S. cerevisiae* DNA (strain S288C, Research Genetics) by PCR using primers Sae2FO (CTCCG TGCTAGCATGGTACTGGTGAAGAAAATG) and Sae2RE (CCAACACCCG GGACATCTAGCATATATCTGC). The PCR product was digested with *NheI* and *XmaI* restriction endonucleases and cloned into *NheI* and *XmaI* sites in pFB-MBP-Sgs1-his<sup>30</sup>, generating pFB-MBP-Sae2-his. Bacmids, primary and secondary viruses were prepared according to manufacturers' recommendations (Bac-to-Bac, Life Technologies). For large-scale infection, the Sf9 cells were seeded at  $0.5 \times 10^6$  cells per ml. The cells were infected the next morning with high-titre virus. The cells were harvested 52 h after infection (500g, 15 min), washed with phosphate buffered saline, snap frozen in liquid nitrogen, and stored at  $-80^\circ\text{C}$ . All subsequent purification steps were carried out at  $0-4^\circ\text{C}$ . This protocol describes purification from 1.61 Sf9 cells. The pellets were thawed, and resuspended in lysis buffer (Tris-HCl, pH 7.5, 50 mM; dithiothreitol, 1 mM; EDTA, 1 mM; Sigma protease inhibitory cocktail, P8340, 1:400; phenylmethylsulphonyl fluoride, 1 mM; leupeptin,  $30 \mu\text{g ml}^{-1}$ ) up to the total volume of 72 ml. Cells were allowed to swell for 20 min with gentle agitation. 36 ml of 50% glycerol was added to the sample. 7 ml of 5 M NaCl was added slowly (while mixing) to the sample, and incubated for 30 min with gentle agitation. The cell suspension was centrifuged at 55,000g for 30 min to obtain soluble extract. 8 ml of pre-equilibrated amylose resin (New England Biolabs) was added to the supernatant, and batch-incubated for 1 h with gentle agitation. The resin was washed  $5\times$  with 40 ml wash buffer (Tris-HCl, pH 7.5, 50 mM;  $\beta$ -mercaptoethanol, 5 mM; NaCl, 1 M; phenylmethylsulphonyl fluoride, 1 mM; leupeptin,  $10 \mu\text{g ml}^{-1}$ ; glycerol, 10%) batch-wise, and then extensively on a disposable column (Thermo Scientific). MBP-Sae2-His was eluted with elution buffer (Tris-HCl, pH 7.5, 50 mM;  $\beta$ -mercaptoethanol, 5 mM; NaCl, 0.3 M; phenylmethylsulphonyl fluoride, 1 mM; leupeptin,  $20 \mu\text{g ml}^{-1}$ ; glycerol, 10%; maltose, 10 mM). The eluate was treated with PreScission protease ( $12 \mu\text{g}$  of protease per  $100 \mu\text{g}$  MBP-Sae2-His), and the sample was incubated for 3 h at  $4^\circ\text{C}$ . Next, imidazole was added to the sample (10 mM final concentration), followed by 1 ml of pre-equilibrated Ni-NTA agarose (Qiagen). The sample was batch-incubated for 1 h. The resin was washed on a disposable column (Thermo Scientific) with NTA buffer A1 (Tris-HCl, pH 7.5, 50 mM;  $\beta$ -mercaptoethanol, 5 mM; NaCl, 1 M; phenylmethylsulphonyl fluoride, 1 mM; leupeptin,  $20 \mu\text{g ml}^{-1}$ ; glycerol, 10%; imidazole, 58 mM). The resin was then washed with NTA buffer A2 (Tris-HCl, pH 7.5, 50 mM;  $\beta$ -mercaptoethanol, 5 mM; NaCl, 150 mM; phenylmethylsulphonyl fluoride, 1 mM; leupeptin,  $20 \mu\text{g ml}^{-1}$ ; glycerol, 10%; imidazole, 58 mM). Sae2-His was eluted with 0.5 ml fractions of buffer B (Tris-HCl, pH 7.5, 50 mM;  $\beta$ -mercaptoethanol, 5 mM; NaCl, 100 mM; phenylmethylsulphonyl fluoride, 1 mM; leupeptin,  $20 \mu\text{g ml}^{-1}$ ; glycerol, 10%; imidazole, 400 mM). Fractions containing protein were pooled, dialysed 1 h against 1 l dialysis buffer (Tris-HCl, pH 7.5, 50 mM;  $\beta$ -mercaptoethanol, 5 mM; NaCl, 100 mM; phenylmethylsulphonyl fluoride, 0.5 mM; glycerol, 10%), aliquoted, frozen in liquid nitrogen, and stored at  $-80^\circ\text{C}$ . Protein concentrations were estimated using the Bradford method with bovine serum albumin as the protein standard. Typical concentration of Sae2 preparation was  $10-15 \mu\text{M}$  and yield up to  $\sim 3 \text{ mg}$  from 1.61 of Sf9 cells. To prepare the Sae2 variants, the *SAE2* gene in the pFB-MBP-Sae2-his plasmid was mutated using QuikChange II XL site-directed mutagenesis kit (Agilent). All variants were then expressed in Sf9 cells (800 ml), and purified as above.

Recombinant MRX was prepared as described previously<sup>31</sup>; expression vectors were a gift from T. Paull and P. Sung. Briefly, Sf9 cells were infected with an optimized ratio of baculoviruses expressing Mre11-His, Xrs2-Flag and Rad50 factors. Proteins were extracted and soluble extract was obtained as described above for Sae2. Recombinant MRX was purified as a complex by affinity chromatography with Ni-NTA agarose (Qiagen) and anti-Flag affinity resin (Sigma, A2220). Recombinant Mre11-Xrs2 complex was prepared in the same way using Mre11-His and Xrs2-Flag constructs. Recombinant Mre11 was purified using affinity (Ni-NTA agarose) and ion exchange (HiTrap Q, GE Healthcare) chromatography. The same procedure we used for WT MRX was unsuccessful when we attempted to purify the nuclease-dead Mre11(H125L;D126V)-Rad50-Xrs2 variant, referred to as MRX-nd for simplicity. The MRX-nd complex was falling apart during the washing steps on Ni-NTA resin. To address this, soluble extract was prepared with only 0.5 mM  $\beta$ -mercaptoethanol and directly applied on the anti-Flag affinity resin (Sigma). The bound complex was washed extensively with wash buffer (Tris-HCl, pH 7.5, 33 mM; EDTA, 0.7 mM; phenylmethylsulphonyl fluoride, 0.5 mM;  $\beta$ -mercaptoethanol, 0.33 mM; glycerol, 16.6%; NaCl, 300 mM; NP40, 0.1%), followed by washing with the same buffer but without NP40, and eluted with Flag peptide (Sigma, Extended Data Fig. 3a). The same procedure was also used for the preparation of wild-type MRX (Extended Data Fig. 3a) as well as MR(K40A)X, MR(K40R)X and MR(K81I)X variants (Extended Data Fig. 3c). Throughout the manuscript, mutant MRX complexes were always compared with the wild-type MRX purified using an identical procedure. The wild-type MRX purified using the second procedure showed  $\sim$ twofold lower specific activity; when double-concentration was used in reactions, its activity

was indistinguishable from the complex prepared using the original procedure<sup>31</sup>. Streptavidin and avidin were purchased from Sigma. Exol<sup>31</sup>, Dna2<sup>32</sup> and RPA<sup>31</sup> were prepared as described previously.

**DNA substrates.** All oligonucleotides were purified on polyacrylamide gels and purchased from Microsynth (Switzerland). The labelling of oligonucleotides at the 5' end was carried out with T4 polynucleotide kinase (New England Biolabs) and [ $\gamma$ -<sup>32</sup>P]ATP. The labelling of oligonucleotides at the 3' end was carried out with terminal deoxynucleotidyl transferase (New England Biolabs) and [ $\alpha$ -<sup>32</sup>P]cordycepin 5' triphosphate. The oligonucleotides used for the 50-bp DNA substrate were PC1253C (AACGTCATAGACGATTACATTGCTAGGACATCTTTGCCACG TTGACCCA) and PC1253B (TGGGTCAACGTGGGCAAGATGTCCTAGCA ATGTAATCGTCTATGACGTT) with 3'-terminal biotin. The oligonucleotides used for the 50-bp dsDNA without biotin label were X12-3 and X12-4C, and for the Y-structure substrate X13-3 and X12-4NC, as described<sup>30</sup>. The oligonucleotides used for the 70-bp DNA substrate were 210 (GTAAGTCCCGCGGTGCGGGTGC CAGGGCGTGCCTTGGGTCTCCCGGCGCGTACTCCACCTCATGCATC) and 211 (GATGCATGAGGTGGAGTACGCGCCCGGGGAGCCCAAGGGCAC GCCCTGGCACCCGACCGCGGCACCTTAC). Internal thymidine (T, in bold) contained biotin label, where indicated. The oligonucleotides used to prepare the 100-mer were 100TOP (GTAAGTCCCGCGGTGCGGGTGCCAGGGCGTGCC CTTGGGCTCCCGGGGCGGTACTCCACCTCATAATCTTCTGCCATGGTC GTAGCAGCCTCTGTCATC) and 100BOTTOM (GATGCAGGAGGCTGCTA CGACCATGGCAGAAGATTATGAGGTGGAGTACGCGCCCGGGGAGCCCA AAGGGCAGCCCTGGCACCCGACCGCGGCACCTTAC). The sequence of HP-2 DNA was described previously<sup>4</sup>, and HL-3 was (ATCATTGCCTATCCTGA CAGTCCGACACACATCGGACTGTCAGGATAGGCAATGATCTTTTT TTT). To prepare the 2.7-kb-long dsDNA substrate with biotin labels used in Extended Data Fig. 8a, b, pATTP-S vector was first prepared by annealing self-complementary oligonucleotide 202 (AGCTGTAGTGCCTCACTGGGGTAA CCTTTGAGTTCTCTCAGTTGGGGGCGTAG) with itself, and cloning it into a HindIII site of pUC19. Then, <sup>32</sup>P-labelled oligonucleotides (variants of 210 and 211 oligonucleotides with the desired biotin modifications) were annealed and reacted in eightfold excess over pATTP-S vector with  $\Phi$ C31 integrase in a linking buffer (Tris-acetate, pH 7.5, 10 mM; EDTA, 1 mM; NaCl, 0.1 M; dithiothreitol, 5 mM; BSA, 10 mg ml<sup>-1</sup>). The resultant linear DNA containing the desired modifications at both ends was separated and purified from agarose gels. M13 ssDNA was purchased from New England Biolabs.

**Nuclease assays.** Unless indicated otherwise, nuclease assays were carried out in 25 mM Tris-acetate, pH 7.5, 1 mM dithiothreitol, 5 mM magnesium acetate, 1 mM manganese acetate, 1 mM ATP, 80 U ml<sup>-1</sup> pyruvate kinase (Sigma), 1 mM phosphoenolpyruvate, 0.25 mg ml<sup>-1</sup> bovine serum albumin (New England Biolabs) and 1 nM (in molecules) DNA substrates. Where indicated, the reactions were supplemented with streptavidin or avidin (Sigma,  $\sim$ 15-fold excess over biotin labels, 15–30 nM) and pre-incubated for 5 min at room temperature. Purified proteins were then added on ice. The reactions were incubated for 30 min at  $30^\circ\text{C}$  and analysed on 15% denaturing polyacrylamide gels (acrylamide:bisacrylamide, 19:1, Bio-Rad), unless indicated otherwise. The gels were fixed in a solution containing 40% methanol, 10% acetic acid and 5% glycerol for 30 min, dried on DE81 chromatography paper (Whatman), and exposed to storage phosphor screens (GE Healthcare). The screens were scanned by a Typhoon phosphor imager (GE Healthcare).

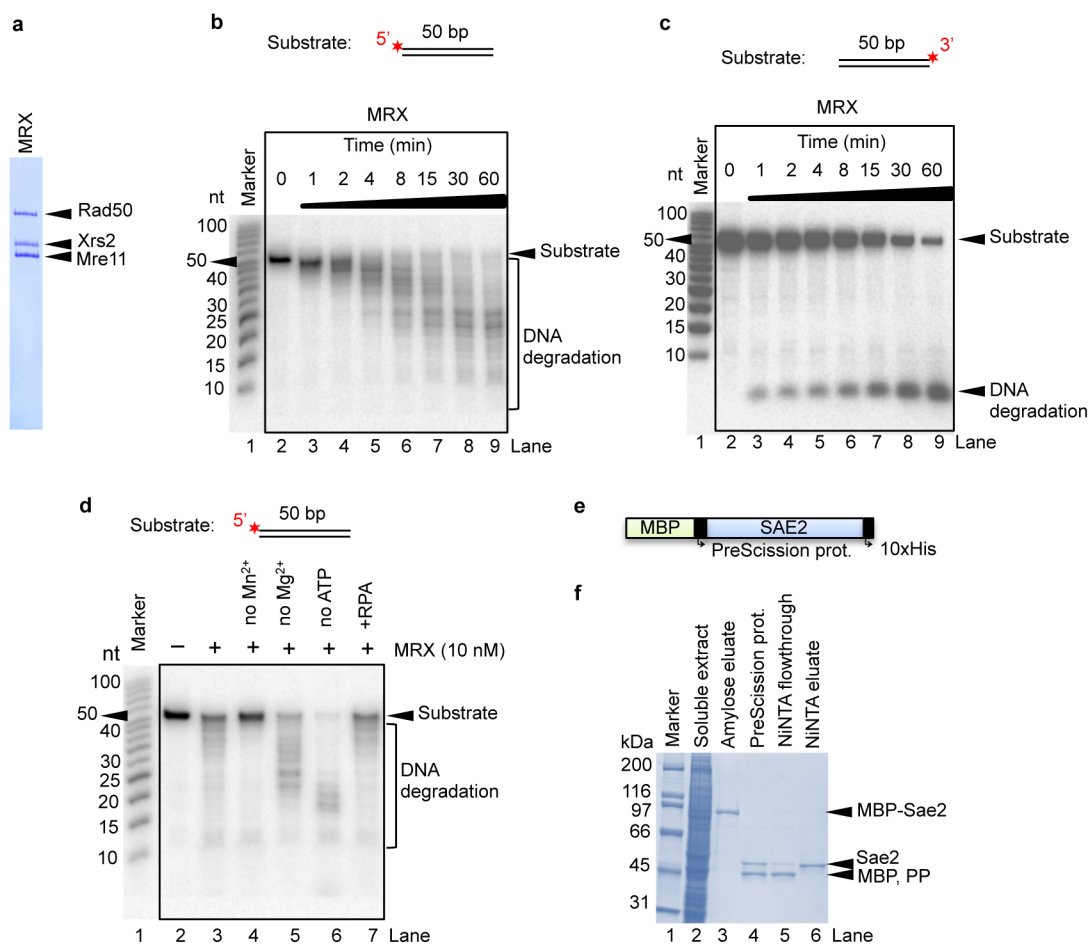
**Electrophoretic mobility shift assay.** The DNA binding capacity of Sae2 mutants was estimated by electrophoretic mobility shift assay, as described previously<sup>30</sup>. The substrate was 270-bp-long DNA generated by PCR using <sup>32</sup>P-labelled primers 224 (GGCCGTCGTTTTACAACGTCGT) and 237 (GGTCGGGGCTGGCTTA ACTATG) and pUC19 dsDNA as the template.

**Protein interaction studies.** To test for interactions between Sae2 and MRX, MBP-Sae2 was expressed in Sf9 cells, cells were lysed, and  $\sim 4 \mu\text{g}$  MBP-Sae2 was bound to amylose resin (50  $\mu\text{l}$ ). The resin was washed with wash buffer (Tris-HCl, pH 7.5, 50 mM; EDTA, 2 mM; NaCl, 80 mM; NP40, 0.2%) and incubated for 1 h at  $4^\circ\text{C}$  with recombinant purified MRX (4  $\mu\text{g}$ ). The resin with bound proteins was then extensively washed with wash buffer, and proteins were eluted with wash buffer (100  $\mu\text{l}$ ) containing 20 mM maltose. The proteins in the eluate were analysed by SDS-PAGE stained with silver or by western blotting using anti-Flag primary antibody (Sigma, F3165) against Xrs2-Flag using standard procedures. Where indicated, proteins in the eluate were treated with PreScission protease to distinguish MBP-Sae2 from Mre11 (tagged Sae2 co-migrates with Mre11, PreScission protease cleaves MBP tag off Sae2).

- Cejka, P. & Kowalczykowski, S. C. The full-length *Saccharomyces cerevisiae* Sgs1 protein is a vigorous DNA helicase that preferentially unwinds Holliday junctions. *J. Biol. Chem.* **285**, 8290–8301 (2010).
- Cannavo, E., Cejka, P. & Kowalczykowski, S. C. Relationship of DNA degradation by *Saccharomyces cerevisiae* exonuclease 1 and its stimulation by RPA and

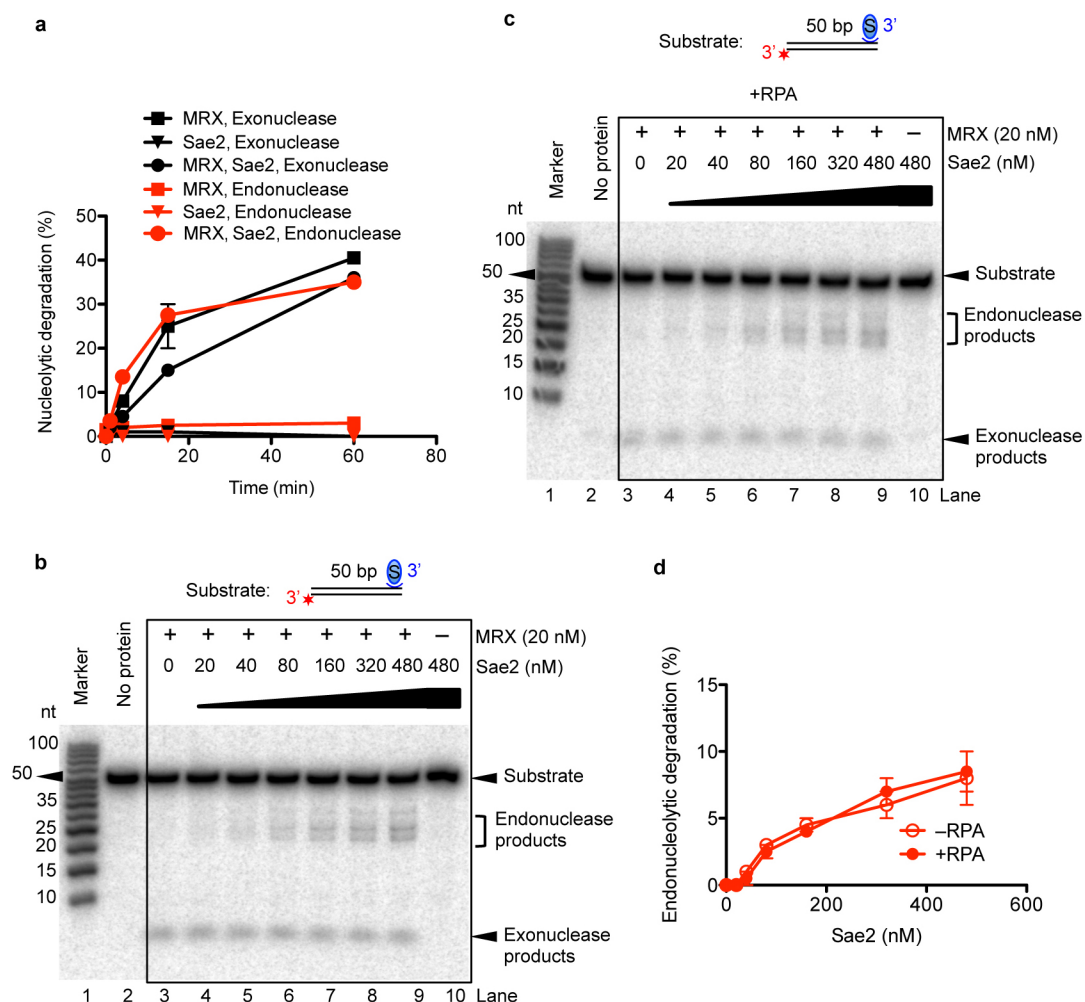
- Mre11-Rad50-Xrs2 to DNA end resection. *Proc. Natl Acad. Sci. USA* **110**, E1661–E1668 (2013).
32. Levikova, M., Klaue, D., Seidel, R. & Cejka, P. Nuclease activity of *Saccharomyces cerevisiae* Dna2 inhibits its potent DNA helicase activity. *Proc. Natl Acad. Sci. USA* **110**, E1992–E2001 (2013).
33. Nicolette, M. L. *et al.* Mre11-Rad50-Xrs2 and Sae2 promote 5' strand resection of DNA double-strand breaks. *Nature Struct. Mol. Biol.* **17**, 1478–1485 (2010).
34. Lisby, M., Barlow, J. H., Burgess, R. C. & Rothstein, R. Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. *Cell* **118**, 699–713 (2004).





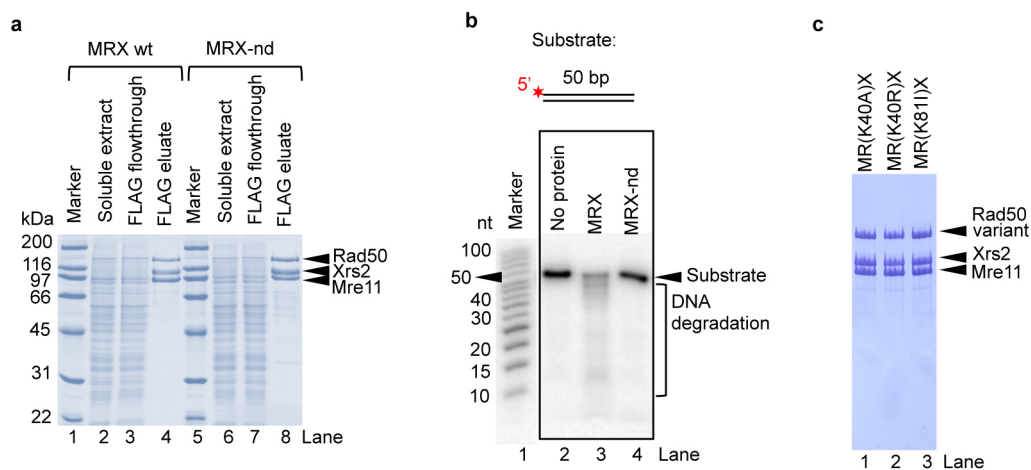
**Extended Data Figure 1 | Purification of wild-type Mre11-Rad50-Xrs2 (MRX) and Sae2.** **a**, Purified MRX used in this study. Gel was stained with Coomassie brilliant blue. **b**, Nuclease activity of MRX (10 nM) on 5'-labelled dsDNA substrate. Products were separated on a denaturing gel. MRX gradually shortened dsDNA with a 5' <sup>32</sup>P label. **c**, Nuclease activity of MRX (10 nM) on 3'-labelled dsDNA substrate. MRX directly released the radioactive label from a 3' <sup>32</sup>P-labelled DNA substrate, showing that it is a 3'-5' exonuclease<sup>4</sup>.

**d**, Nuclease activity of MRX and its dependence on manganese, ATP and RPA, as indicated. The MRX exonuclease requires manganese (5 mM), is moderately inhibited by ATP (1 mM), and is not affected by RPA (23 nM). **e**, A scheme of the Sae2 construct. Sae2 contains an N-terminal MBP tag and a C-terminal His-tag (10×His). **f**, Representative purification of Sae2. Gel was stained with Coomassie brilliant blue. MBP affinity tag was cleaved off during protein purification. MBP, maltose-binding protein; PP, PreScission protease.



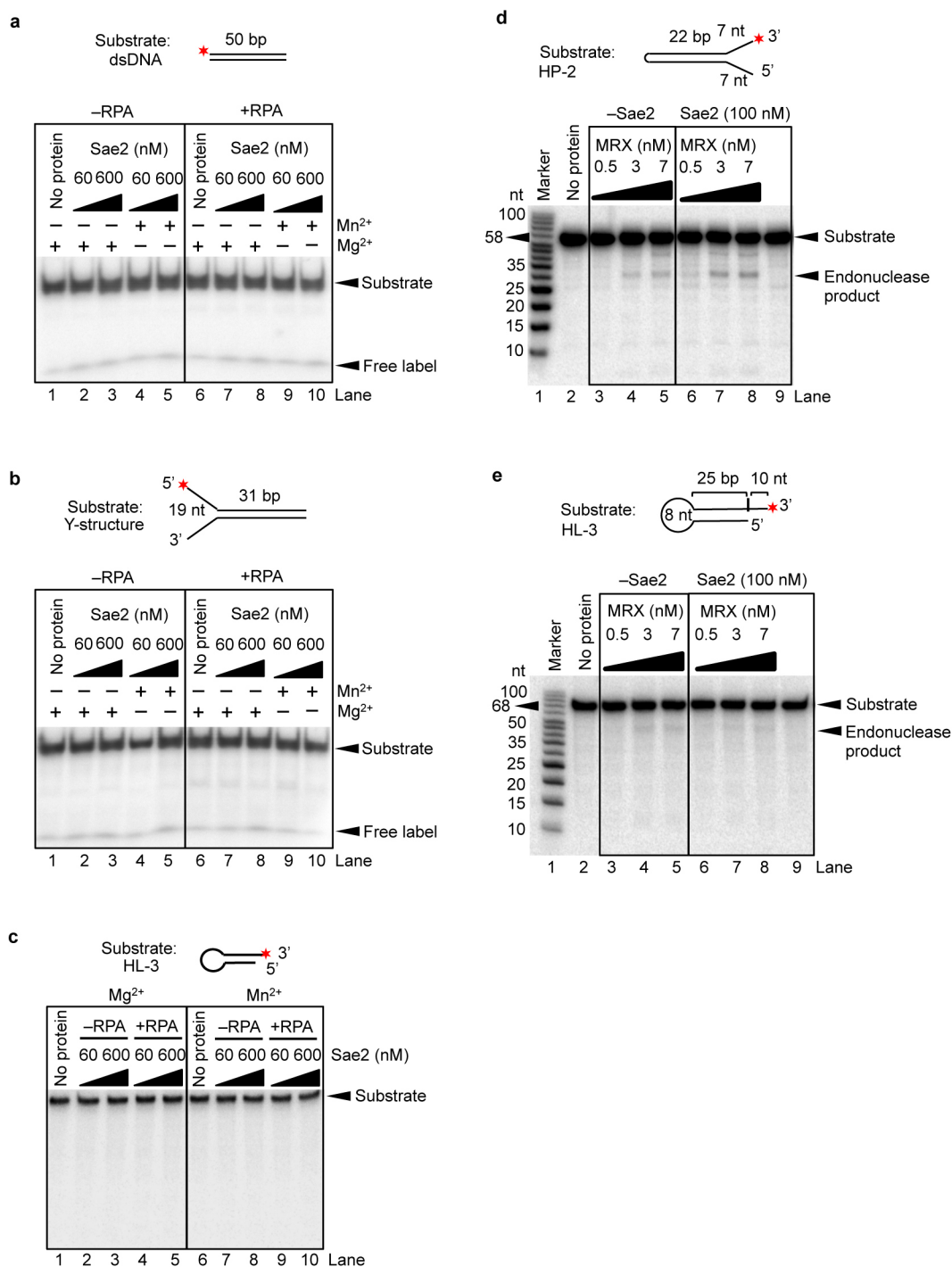
**Extended Data Figure 2 | MRX dsDNA endonuclease is promoted by Sae2 and unaffected by RPA.** **a**, Quantitation of the data such as from Fig. 1a. Averages shown,  $n = 2$ , error, s.e.m. **b**, Nuclease assay (15 min) was carried out with MRX (20 nM) and a range of Sae2 concentrations, as indicated. The exonuclease of MRX was unaffected by Sae2, but endonuclease cleavage

increased with Sae2 concentration. **c**, Nuclease activity of MRX and Sae2 on dsDNA is not affected by the single-strand DNA binding protein RPA. Nuclease assay was carried out as in **b**, but with RPA (23 nM). **d**, Quantitation of the data such as from **b** and **c**. Averages shown,  $n = 2$ ; error bars, s.e.m.



**Extended Data Figure 3 | Purification of MRX variants.** **a**, Polyacrylamide gel electrophoresis showing representative purifications. Gel was stained with Coomassie brilliant blue. **b**, Nuclease assay was performed with 5'-labelled dsDNA substrate and either wild-type or nuclease-deficient M(H125L;D126V)RX variant (MRX-nd). As expected, the mutant MRX

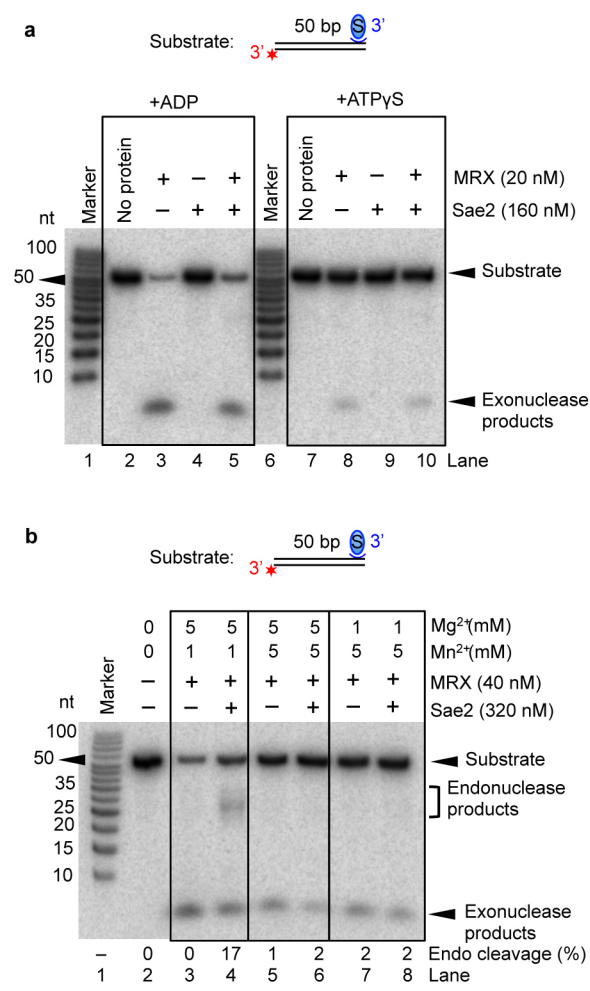
possessed no activity, indicating that the nuclease is intrinsic to Mre11. **c**, Purified MRX mutants used in this study. MR(K40A)X is expected not to bind ATP; MR(K40R)X is expected to bind but not to hydrolyse ATP; MR(K81I)X is a Rad50S MRX variant.



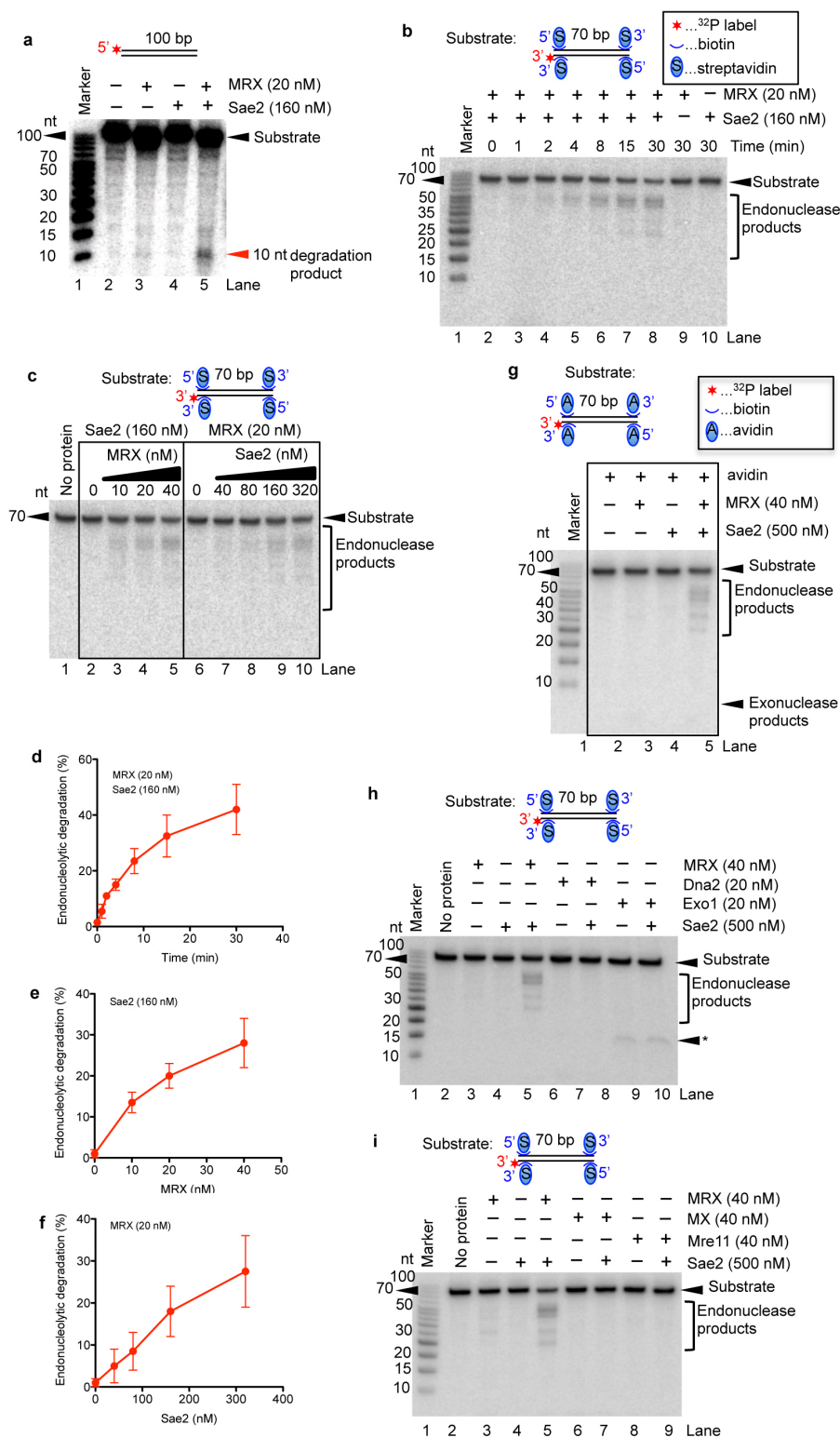
**Extended Data Figure 4 | Sae2 does not show nuclease activity, and does not promote MRX nuclease on hairpins.** **a**, Recombinant Sae2 was assayed on a dsDNA substrate in the presence of either magnesium (5 mM) or manganese (5 mM), with or without RPA (23 nM), as indicated. Free label, carryover of [<sup>32</sup>P]ATP from the labelling reaction, marks the position of the smallest possible product resulting from potential nuclease activity. Samples were analysed on a 10% native polyacrylamide gel. **b**, Recombinant Sae2 was assayed on a Y-structure DNA substrate either in the presence of magnesium (5 mM) or manganese (5 mM), with or without RPA (23 nM), as indicated. Free label, carryover of [<sup>32</sup>P]ATP from the labelling reaction, marks the position of the smallest possible product resulting from potential nuclease activity. Samples

were analysed on a 10% native polyacrylamide gel. **c**, Recombinant Sae2 was assayed on a HL-3 hairpin DNA substrate either in the presence of magnesium (5 mM) or manganese (5 mM), with or without RPA (23 nM), as indicated. Samples were analysed on a 15% denaturing polyacrylamide gel. **d**, Nuclease assay was performed with MRX and Sae2 on HP-2 DNA, as indicated. Samples were analysed on a 15% denaturing polyacrylamide gel. Sae2 does not promote endonuclease of MRX on HP-2 DNA substrate. **e**, Nuclease assay was performed with MRX and Sae2 on HL-3 hairpin DNA, as indicated. Samples were analysed on a 15% denaturing polyacrylamide gel. Sae2 does not promote endonuclease of MRX on HL-3 DNA substrate.





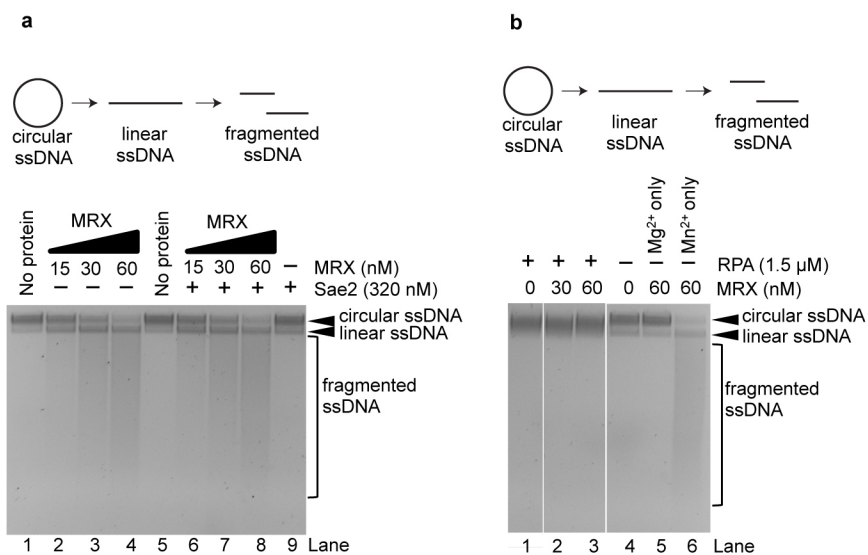
**Extended Data Figure 5 | Analysis of Sae2 and MRX endonuclease activity on 50-bp single-blocked dsDNA substrate.** **a**, Nuclease assay was performed with ADP or non-hydrolysable ATP analogue ATPγS, and MRX and/or Sae2, as indicated. Neither ADP nor ATPγS supported the endonuclease activity, suggesting that ATP hydrolysis is essential. **b**, Nuclease assay was performed with MRX and/or Sae2 and various concentrations of magnesium and manganese, as indicated. Higher concentration of magnesium than manganese is required for the endonuclease of MRX stimulated by Sae2. Endo cleavage (%), average percentage of endonucleolytic products, on the basis of two independent experiments.



### Extended Data Figure 6 | Analysis of Sae2 and MRX-dependent endonuclease activity.

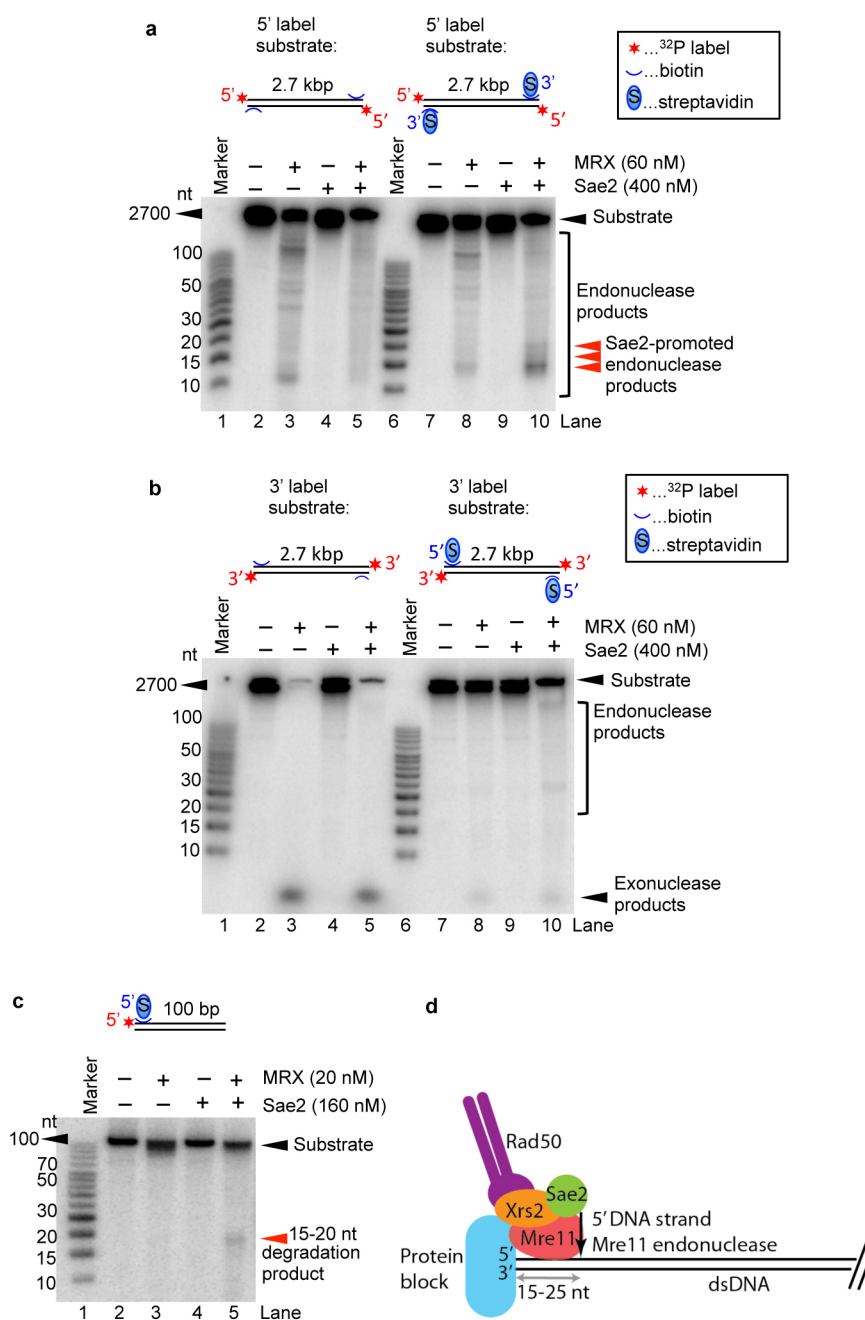
**a**, Nuclease assay was performed with a 5'-labelled 100-bp-long dsDNA substrate in the absence of a protein block. MRX and Sae2 concentrations were used as indicated. Sae2 promoted MRX endonuclease even in the absence of a protein block; however, the reaction was inefficient and contrast of the image had to be enhanced to visualize the degradation products. The cleavage occurred ~10 nucleotides away from the DNA end, which is different from protein-blocked substrates, which were cleaved typically ~15–20 nucleotides away from the end. **b**, Kinetic analysis of MRX and Sae2 endonuclease activity. The preferred position of cleavage is located ~50 nucleotides from the 3' end, and ~20 nucleotides from the 5' end. **c**, Nuclease

assay (15 min) was performed with indicated concentrations of MRX and Sae2. The extent of endonuclease cleavage is dependent on concentrations of both Sae2 and MRX. **d**, Quantification of the data from **b**. Averages shown,  $n = 2$ , error bars, s.e.m. **e**, **f**, Quantification of the data from **c**. Averages shown,  $n = 2$ , error bars, s.e.m. **g**, Experiment as in Fig. 2, but with avidin instead of streptavidin. Both avidin and streptavidin promote Sae2 and MRX endonuclease to a similar extent, showing that there is no need for a specific interaction between Sae2–MRX and the protein block. **h**, **i**, Nuclease assays were performed with recombinant proteins as indicated. MX, Mre11–Xrs2. Sae2 promotes only the endonuclease of MRX. \*, Nucleolytic product resulting from Exo1 activity, independent of Sae2.



**Extended Data Figure 7 | Sae2 does not promote ssDNA endonuclease activity of MRX.** **a**, Circular M13 ssDNA was used as a substrate in a nuclease assay with MRX and/or Sae2, as indicated. Sae2 did not affect the ssDNA endonuclease of MRX. Reaction products were analysed on 1% agarose gel and stained with Gel red (Invitrogen). **b**, The ssDNA endonuclease activity of MRX

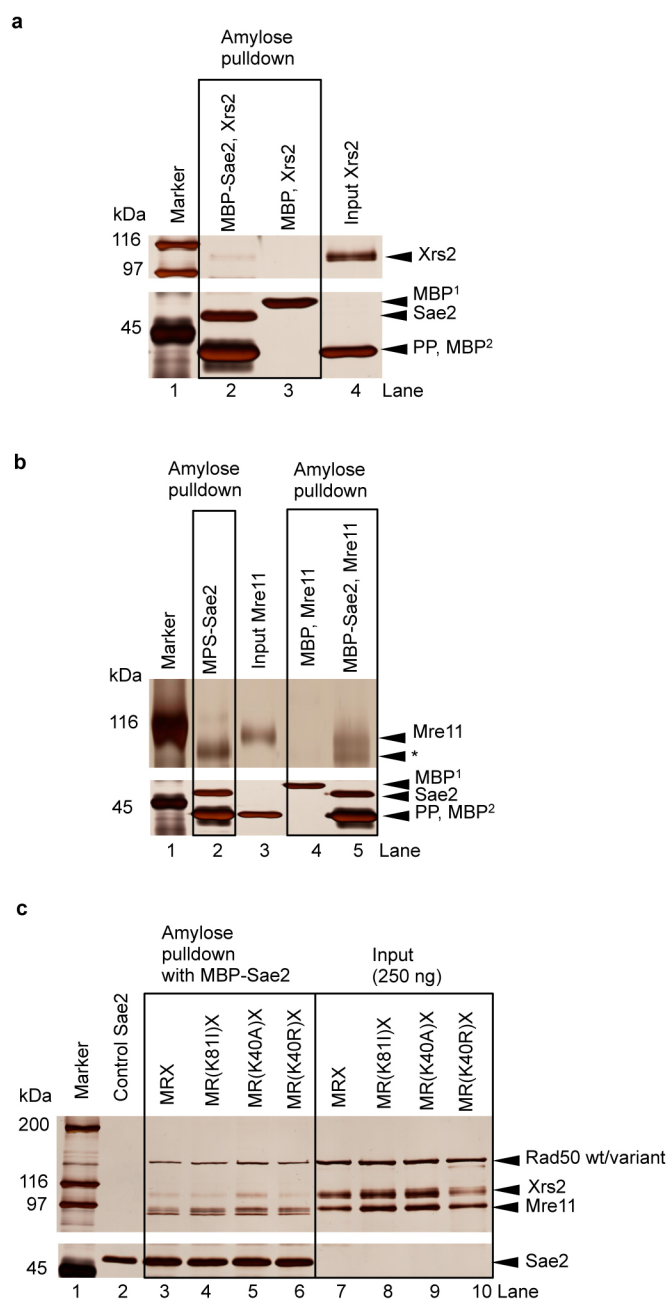
is dependent on manganese and inhibited by saturating concentrations of RPA (1.5 μM). Mg<sup>2+</sup> only, 5 mM magnesium acetate, no manganese; Mn<sup>2+</sup> only, 5 mM manganese acetate, no magnesium. Reaction products were analysed on 1% agarose gel and stained with Gel red.



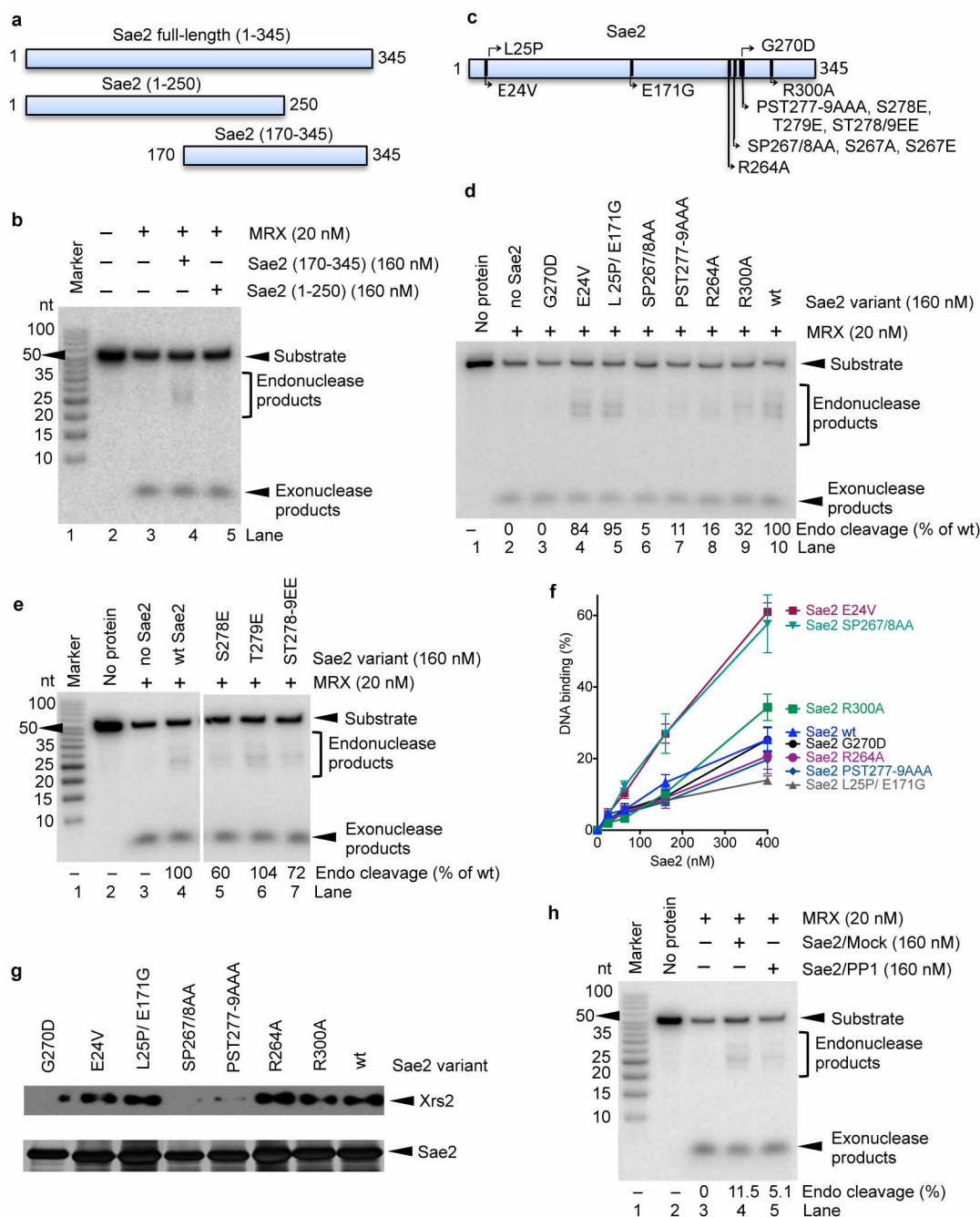
**Extended Data Figure 8 | Sae2 and MRX endonuclease preferentially cleave 5'-terminated DNA strand.** **a**, Nuclease assay was performed with a 5' <sup>32</sup>P-labelled 2.7-kb-long dsDNA substrate either without streptavidin (lanes 2–5) or with streptavidin (lanes 7–10), and MRX and Sae2, as indicated, for 60 min. The 2.7-kb-long substrate was prepared by reacting pATTP-S plasmid with annealed labelled oligonucleotides and ΦC31 integrase, as described in Methods. Reaction products were separated on a 15% denaturing polyacrylamide gel. Unprocessed DNA substrate did not enter the gel and remained trapped in the wells. MRX alone has the capacity to cleave dsDNA endonucleolytically at various distances from the 5' end (lane 3), in agreement with previous reports<sup>24,33</sup>. This endonuclease activity is not affected by the protein block (compare lanes 3 and 8). Sae2 promotes endonucleolytic cleavage

specifically near the protein-blocked DNA end (lane 10, indicated by red arrows). **b**, Assay as in **a**, but with a 3'-labelled DNA substrate. No endonuclease activity of MRX and Sae2 near the 3' end was detected. **c**, Nuclease assay as in Fig. 3a, but with a DNA substrate of 100 bp in length (instead of 70 bp). Concerted action of MRX and Sae2 resulted in DNA cleavage ~15–20 nucleotides away from the streptavidin-blocked 5' DNA end. The position of the cleavage was identical for both 100- and 70-bp-long DNA substrates (compare with Fig. 3a), suggesting that the protein-blocked DNA end directs the position of cleavage by MRX and Sae2. **d**, A cartoon depicting the position of endonucleolytic cleavage by MRX–Sae2. For simplicity, the MRX complex is depicted as a monomer.





**Extended Data Figure 9 | Sae2 interacts with Mre11 and Xrs2 subunits of the MRX complex.** **a**, Amylose pull-down was carried out with MBP-Sae2 and Xrs2 or MBP and Xrs2. Xrs2 bound to MBP-Sae2 (lane 2) but not to MBP (lane 3), showing that Xrs2 binds Sae2 but not the MBP tag or the amylose resin. We point out that the interaction was very weak, and the amount of Xrs2 we pulled down with MBP-Sae2 was near the limit of detection by silver staining. Lane 4, 63 ng of recombinant Xrs2 was loaded as a control. Samples in lanes 2 and 4 were treated with PreScission protease before gel analysis. MBP<sup>1</sup>, maltose-binding protein expressed in *E. coli*; MBP<sup>2</sup>, maltose-binding protein resulting from cleavage with PreScission protease; PP, PreScission protease. **b**, Amylose pull-down was carried out with MBP-Sae2 and Mre11 or MBP and Mre11. Mre11 bound to MBP-Sae2 (lane 5), but not to MBP (lane 4), showing that Mre11 binds Sae2 but not the MBP tag or the amylose resin. Lane 3, 10 ng of recombinant Mre11 was loaded as a control. Lane 2, amylose pull-down was carried out with MBP-Sae2 alone (without Mre11). The interaction between Sae2 and Mre11 is very weak, as the amount of Mre11 we pulled down with MBP-Sae2 is very small. Samples in lanes 2, 3 and 5 were treated with PreScission protease. The band migrating just below Mre11 (indicated by an asterisk) is likely to be residual uncleaved MBP-Sae2. The image in the upper panel was stretched vertically for visualization purposes. MBP<sup>1</sup>, maltose-binding protein expressed in *E. coli*; MBP<sup>2</sup>, maltose-binding protein resulting from cleavage with PreScission protease; PP, PreScission protease. **c**, Amylose pull-down was carried out with MBP-Sae2 and variants of MRX (lanes 3–6). Lanes 7–10, 250 ng of the respective MRX preparations was loaded as a control. Control Sae2, 120 ng recombinant Sae2. Sae2 interacts with both MRX variants deficient in ATP binding and/or hydrolysis, as expected. Sae2 also interacts with MR(K81I)X, indicating that the defects in the activation of the endonuclease of the Rad50S MRX variant by Sae2 cannot simply be explained by a lack of interaction, which is in accordance with proper Sae2 recruitment to double strand breaks in *rad50s* mutants<sup>34</sup>. However, since Sae2 likely interacts with multiple subunits of the MRX complex, we cannot exclude a defect in a subset of the interaction sites, which may abrogate the functional interplay between MR(K81I)X and Sae2.



**Extended Data Figure 10 | C-terminal region of Sae2 is critical for the stimulation of the MRX endonuclease.** **a**, A scheme depicting Sae2 truncation mutants analysed in this study. **b**, Nuclease assay (with 50-mer single-blocked DNA) was performed with MRX and N or C-terminal truncation mutants of Sae2, as indicated. N-terminal region of Sae2 is dispensable, while C-terminal region is essential for the stimulation of the MRX endonuclease. **c**, A scheme depicting Sae2 mutants analysed in this study. **d**, **e**, Nuclease assay (with 50-mer single-blocked DNA) was performed with MRX and Sae2 variants, as indicated. Endo cleavage (% of WT), average percentage of endonucleolytic products, normalized to wild-type Sae2, on the basis of two independent experiments. **f**, Electrophoretic mobility shift assay was used to determine the capacity of Sae2 variants to bind DNA. The results (average) are based on disappearance of the substrate band;  $n = 3$ , error bars, s.e.m. **g**, Amylose pull-down was carried

out with MBP-tagged Sae2 variants and MRX. The presence of MRX in the pull-downs was detected by western blotting using anti-Flag antibody against Xrs2. Sae2 was detected by silver staining. **h**, Nuclease assay (with 50-mer single-blocked DNA) was performed with MRX and Sae2 either mock-treated (incubated with protein phosphatase 1 reaction buffer for 15 min at 30 °C) or protein phosphatase 1-treated Sae2 (New England Biolabs, 1.25 U per 2.5  $\mu$ g recombinant Sae2, 15 min at 30 °C). Endo cleavage (%), average percentage of endonucleolytic products, on the basis of three independent experiments. Treatment of Sae2 with protein phosphatase 1 leads to a reduction of Sae2 capacity to promote MRX endonucleolytic activity. This suggests that Sae2 purified from Sf9 cells is phosphorylated, and this post-translational modification promotes its capacity to stimulate MRX endonuclease.

## CORRIGENDUM

doi:10.1038/nature13719

### **Corrigendum: Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators**

Daniel W. Bellott, Jennifer F. Hughes, Helen Skaletsky, Laura G. Brown, Tatyana Pyntikova, Ting-Jan Cho, Natalia Koutseva, Sara Zaghlul, Tina Graves, Susie Rock, Colin Kremitzki, Robert S. Fulton, Shannon Dugan, Yan Ding, Donna Morton, Ziad Khan, Lora Lewis, Christian Buhay, Qiaoyan Wang, Jennifer Watt, Michael Holder, Sandy Lee, Lynne Nazareth, Jessica Alföldi, Steve Rozen, Donna M. Muzny, Wesley C. Warren, Richard A. Gibbs, Richard K. Wilson & David C. Page

*Nature* **508**, 494–499 (2014); doi:10.1038/nature13206

Jessica Alföldi should have been listed with affiliation 1 in the author list. She performed BAC mapping, radiation hybrid mapping and real-time polymerase chain reaction analyses. The online versions of this Article have been corrected.

## CORRIGENDUM

doi:10.1038/nature13843

### Corrigendum: Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells

Ryan Lister, Mattia Pelizzola, Yasuyuki S. Kida, R. David Hawkins, Joseph R. Nery, Gary Hon, Jessica Antosiewicz-Bourget, Ronan O'Malley, Rosa Castanon, Sarit Klugman, Michael Downes, Ruth Yu, Ron Stewart, Bing Ren, James A. Thomson, Ronald M. Evans & Joseph R. Ecker

*Nature* **471**, 68–73 (2011); doi:10.1038/nature09798

The parameters described in the “Identification of DMRs” subsection of the Methods of our Article regarding the identification of the non-CG mega-DMRs (differentially methylated regions) should read as follows: “The average methylation level of mC called (1% FDR) in the mCHG sequence context was determined in 1-kb windows (sW). The genome was scanned considering groups of 50 adjacent windows sW. The set of 50 average values in the H1 sample was compared to the set of 50 average values in the ADS-iPSC sample using the Wilcoxon test.” We thank Mark van de Wiel for bringing this to our attention. Importantly, the specific code used for this analysis can be found in the methylPipe R package on the Bioconductor website (<http://bioconductor.org/>).

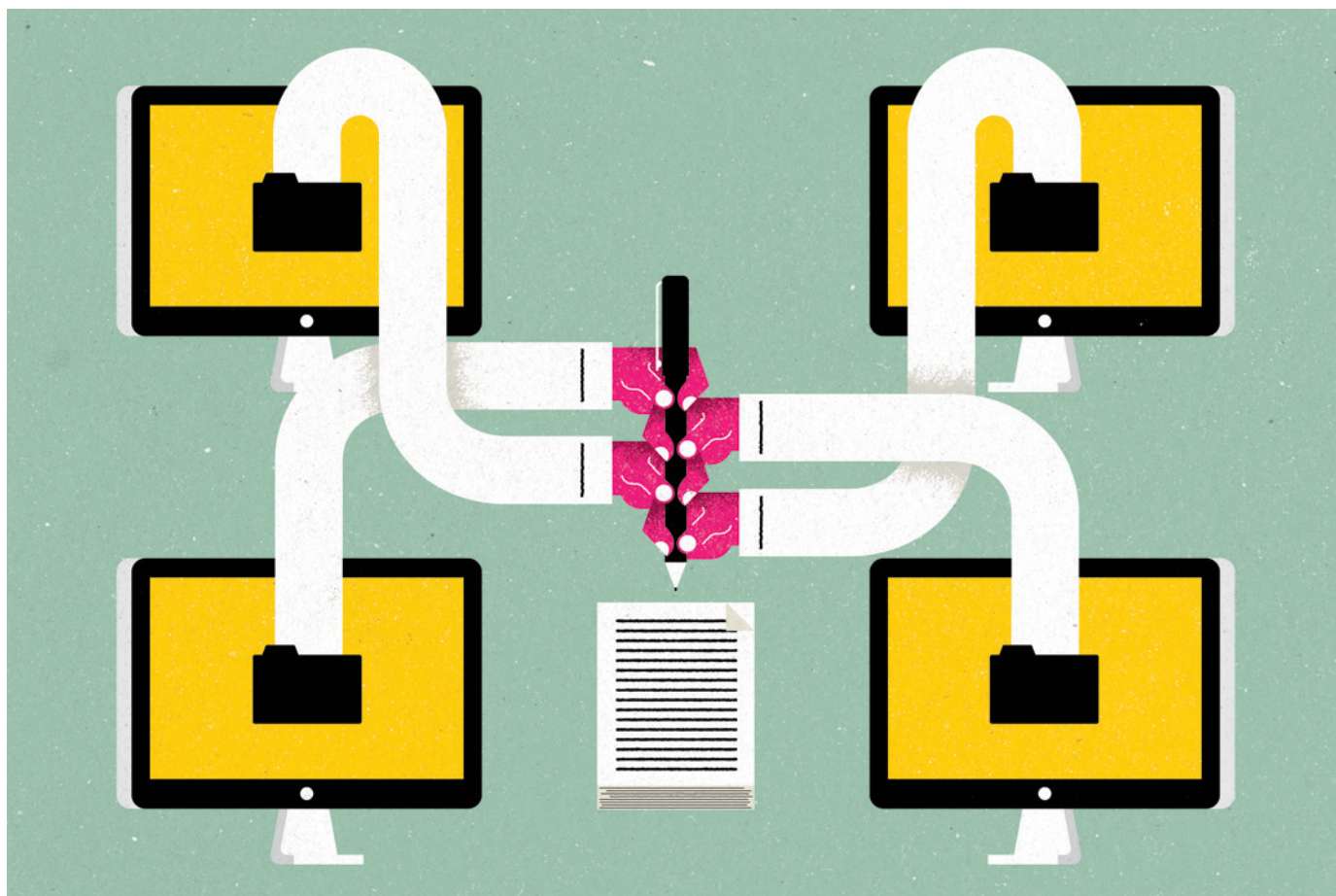


## TOOLBOX

# SCIENTIFIC WRITING: THE ONLINE COOPERATIVE

*Collaborative browser-based tools aim to change the way researchers write and publish their papers.*

ILLUSTRATION BY THE PROJECT TWINS



BY JEFFREY M. PERKEL

When Fernando Cagua was preparing to write up his findings on the economics of whale-shark tourism, he didn't fire up Microsoft Word. He opened his web browser.

Cagua, an ecologist at King Abdullah University of Science and Technology in Thuwal, Saudi Arabia, was keen to try out an online writing environment that would allow him and his three co-authors to work on the same paper simultaneously. Over the past few years, a small cadre of tools have sprung up expressly for this purpose.

Although the features vary, each is designed to ease a key difficulty in writing multi-authored research papers: handling collaboration. And some of the creators have wider ambitions — to fundamentally alter the way that scientific papers are written and published.

Writing a paper is traditionally a stepwise process. One author shares drafts with colleagues and then waits for everyone to reply or moves forward independently, folding in revisions and queries as they arrive. The more co-authors, the more complicated this gets, says Russell Neches, a microbiology PhD student at the University of California, Davis. “Man-

aging that process can be more difficult, more time-consuming and more work than the research itself,” he says.

Collaborative tools simplify this process by allowing multiple authors to edit and format an online document at the same time. The most widely used general-purpose collaborative writing app is probably Google Docs — essentially a stripped-down, online version of Microsoft Word. But there are also more-technical tools designed specifically for researchers. These applications add options such as the ability to control a document's layout and to add citations in a way that suits scientific manuscripts. The ►

► tool that Cagua had his eye on, for instance, writeLaTeX, was so named because it uses the typesetting computer language LaTeX — popular among physical scientists and mathematicians for rendering mathematical formulas, tables and figures. (The tool is produced by a company also called writeLaTeX, supported by Digital Science, a division of Macmillan, which publishes *Nature*. In January, the firm relaunched the tool and renamed it Overleaf.) Other scholar-focused online writing apps include shareLaTeX, Fidus Writer and Authorea.

***“We believe in the idea of an interactive, data-driven article”***

## WORD OF MOUTH

A minority of researchers use these apps, but their number is growing. In the past year, registered users of Overleaf have reached 100,000, says writeLaTeX co-founder John Hammersley, and they have created more than 1.4 million documents with the tool. Authorea has 10,000 users, according to its co-founder Alberto Pepe. Jenna Morgan Lang, a postdoc in the same group as Neches, says that she has one Authorea-written paper in preprint and six more in development. “I do love it,” she says, “and I tell everyone who will listen that they should be using it, too.”

At the heart of the collaborative approach is the way the tools keep track of different versions of the same document. Authorea, for example, breaks documents into user-defined, paragraph-sized chunks that only one author can edit at a time, but multiple researchers can work on different sections of a document simultaneously. The system records every change in a document history. “You can go back and understand how a scientific paper evolved from the first word to the last,” says Pepe.

For Authorea, that concept is based on the software-management system Git, used by programmers to keep track of changes on collaborative code-writing projects, and by data scientists to record their analysis workflow. Other tools take different approaches: Google Docs and Fidus Writer allow all users access to the entire file simultaneously, and track changes more or less like Microsoft Word, but Fidus Writer, for example, does not record the detailed history of every single edit (although a user may save time-stamped document versions). Overleaf allows both a version history and a track-changes facility — but the latter is available only to paying subscribers. Although each tool offers a free account, only researchers willing to pay monthly fees (US\$7–12 for Overleaf and \$5–25 for Authorea) can access the advanced features, such as more storage space or private accounts.

The tools are much more than just word processors and collaboration managers, however. Authorea allows users to build and format bibliographies by searching and importing references from PubMed or CrossRef,

or using DOIs (digital object identifiers); Overleaf allows imports from reference managers Zotero and CiteULike. Authorea also enables users to export documents in any of about 40 different journal formats, including those of *Nature*, *Science* and *Proceedings of the National Academy of Sciences*. By recasting the same data through different journal filters, “it’s a bit like Instagram for scientific papers”, Pepe wrote in one blogpost.

At writeLaTeX, Hammersley has ambitions to integrate the writing and publishing of articles even more closely. Users can click a button to transmit their article directly to journal editors; the company currently has arrangements with around a dozen journals, and many more will follow in the next few months, Hammersley says. However, Cagua says that he did not find the process particularly automatic with a paper he transmitted to *PeerJ*; he had to resubmit information in his original LaTeX file that was not automatically picked up by the journal. But Hammersley says that integration with journals is a work in progress. Ultimately, he hopes that a paper’s author and its journal editor might collaborate on the article together in the browser window.

## FAMILIAR GROUND

Cagua also ended up writing most of his whale-shark paper in Google Docs, because his co-authors were not well versed in LaTeX and so found the original writeLaTeX “too intimidating”. A raw LaTeX file — text interspersed with code that tells typesetting programs how to display the prose and figures — can look off-putting to the uninitiated, or just ugly, like reading the HTML source code behind a web browser’s display. In the relaunched version, Overleaf, a rich-text editing environment hides the code and makes writing friendlier for non-experts. Fidus Writer and Authorea also support LaTeX, as well as other computer languages for controlling the display of raw text, including HTML and Markdown.

Authorea’s “fundamental mission”, Pepe says, “is to re-imagine the scientific article”. Conceived to advance the open sharing of scientific research, the program supports software such as IPython notebooks, which allow readers to explore and manipulate the data underlying published figures. “We believe in the idea of an interactive, data-driven article,” Pepe explains — an idea that he has explored in a prototype ‘Paper of the future’ (see [go.nature.com/plgshx](http://go.nature.com/plgshx)). A few journals are cautiously experimenting with interactive graphics and data in their articles, although for the most part, this is still rare.

An Authorea-written document can double as both a readable paper and an online research notebook containing raw data, notes Alyssa Goodman, an astronomer at Harvard University who was Pepe’s postdoctoral adviser when he developed the software. “The part you can read that looks like a paper is the tip of the iceberg that describes everything

underneath,” she explains.

Using that feature, Neches collaborated with two researchers in Michigan who he chatted with on Twitter but has never met in person. Together, they studied whether materials printed with a 3D printer were sterile for use in bacterial culture experiments. Authorea, he says, provided a forum for team members to upload raw data and methods, from which they could co-assemble a manuscript online. “It was very much as though we had created a laboratory in which we worked together,” he says. “It probably would not have happened at all without a tool like Authorea existing.” ■

**Jeffrey M. Perkel** is a writer based in Pocatello, Idaho.

## MORE ONLINE

Q & A



In the ‘My digital toolbox’ series, scientists share the software and tools they find most useful in their research.

### Ecologist Christie Bahlai (pictured):

“The single greatest data management tool I’ve come across in the past year is OpenRefine. It is a fantastic web-based tool that streamlines the process of cleaning up messy data. And it is, to my knowledge, the only tool of its kind with a user-friendly graphical interface.”

Read more at [go.nature.com/zqybzv](http://go.nature.com/zqybzv)

### Computational nuclear engineer Katy Huff:

“The tool that has most powerfully impacted the reproducibility, transparency and robustness of my work is definitely the combination of Git and GitHub. These are version-control systems; the laboratory notebook of scientific computing.”

Read more at [go.nature.com/l4tsiy](http://go.nature.com/l4tsiy)

### Ecologist Ethan White:

“I learned about the IPython notebook in early 2012, and was immediately hooked. The first time I opened one up it was clear that this tool was going to change the way I worked. I’ve been using it for both teaching and research ever since.”

Read more at [go.nature.com/wz4sny](http://go.nature.com/wz4sny)

For more on scientific software, apps and online tools, visit [nature.com/toolbox](http://nature.com/toolbox)

COURTESY OF CHRISTIE BAH LAI

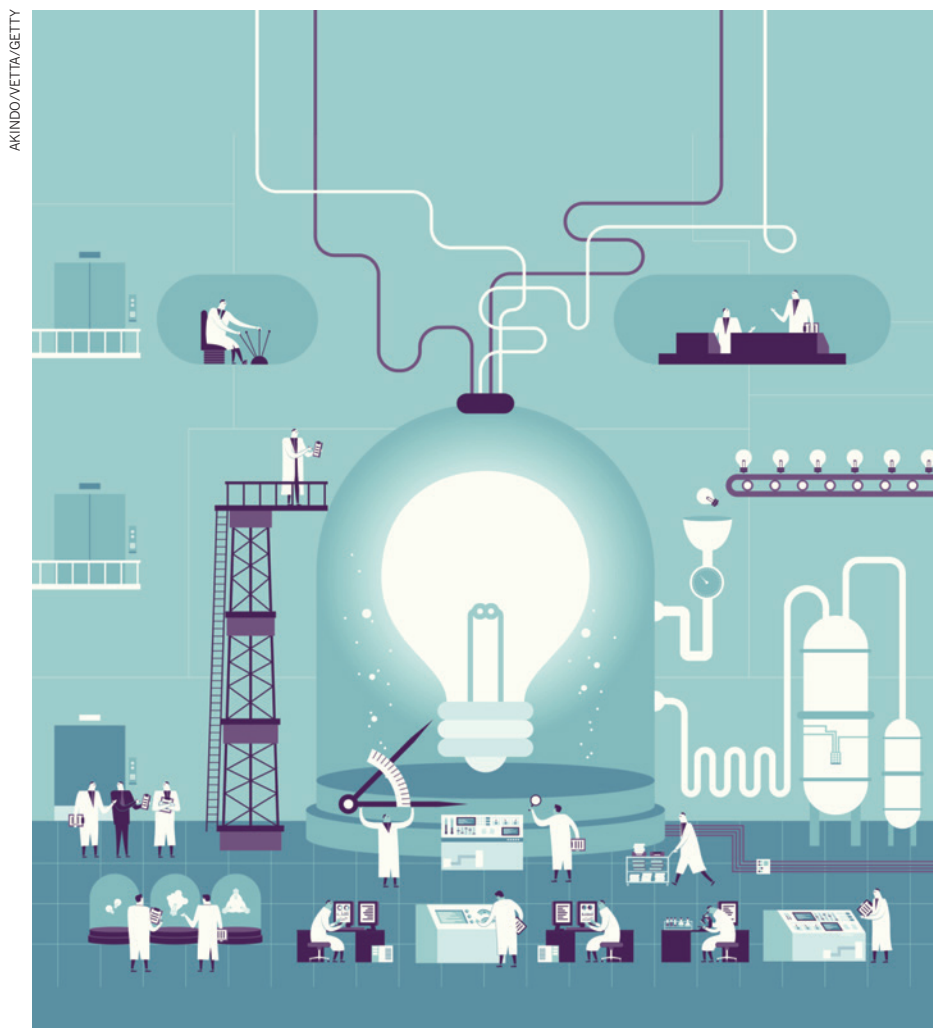


# CAREERS

**TURNING POINT** Colombian postdoc leverages fellowship to fight parasites **p.131**

**NATUREJOBS BLOG** Discussions and tips on careers and jobs [blogs.nature.com/naturejobs](http://blogs.nature.com/naturejobs)

**NATUREJOBS** For the latest career listings and advice [www.naturejobs.com](http://www.naturejobs.com)



AT THE BENCH

## The right mix

*Staffing a lab is fraught with complexity, so new team leaders can learn a lot from the experience of others.*

BY CHRIS WOOLSTON

Evolutionary biologist Erin Kelleher has just started her first lab: she has a technician and would like to bring in a couple of PhD students soon. Chemical and biological engineer Robert Langer oversees an empire of nearly 100 postdocs, graduate students and

technicians. The two are at markedly different career stages — one could hold a lab meeting at a restaurant booth, the other would need an auditorium — but they have something in common. They want each staff member to be just right for their lab — a good worker, a good colleague and, most of all, a good fit.

Of all of the tasks facing lab leaders, staffing

is one of the most important — and the most challenging. Most researchers encounter plenty of cautionary examples as they work their way up through the academic ranks: unfocused graduate students, overwhelmed postdoctoral researchers and surly or sloppy technicians.

Picking the right people is a skill that can take an entire career to perfect. Langer, from the Massachusetts Institute of Technology in Cambridge, has been recruiting staff for more than 30 years, but says he still doesn't think his "interview questions are as good as they could be".

Few principal investigators (PIs) receive instruction in how to staff a lab, and that can open the door for plenty of early-career missteps, says Duncan Odom, a human genetics researcher at the University of Cambridge, UK. "We've done a poor job of training postdocs to become group leaders," he says. "In fact, we haven't really done that job at all. Most postdocs are in large labs that have been running for a long time. They don't have any understanding about what it's like to set up a lab."

### RIGHT ON COURSE

Odom suggests that new PIs take management courses to help them with the transition from researcher to leader. A common mistake, he says, is to quickly add as many workers as a budget will allow. It might be possible to pay their salaries, but a new lab is unlikely to generate enough data, projects and papers to keep everyone happy, engaged and productive.

That leads to turf wars over projects, arguments about authorship and, in some cases, stalled careers. "I've seen labs implode from getting too big too fast," he says. "I deliberately grew my lab slowly. Feeding too many hungry mouths with limited resources is a recipe for trouble." He currently has nine members — two postdocs, two graduate students and five staff scientists, recruiting roughly one each year that his lab has been running. Langer agrees that the pace is important: his lab has roughly tripled in size over the past two decades, but he says he has never suffered from growing pains and has no shortage of projects to go around.

Without careful management, even established labs can become too large for their own good, Odom says. "Most labs with 20 or more people become incubators for Darwinian-type battles, whether they want to or not."

And getting the right people is not an easy job, stresses Frank Chan, who ►

► studies genetics and evolution at the Friedrich Miescher Laboratory, a research institute of the Max Planck Society in Tübingen, Germany. He generally has six or seven lab members at a time — and a lot of other people who would like to be there. With so much interest, he can afford to be discerning. “Of 100 applications, five to ten will be really good,” he says. “It’s a tough market for both sides. It’s hard to find a match.”

Chan attracts applicants from all over the world, which means that in-person interviews are rarely an option. Still, he always talks to potential lab members either on the phone or, even better, over Skype. He wants people who have a solid, career-based reason for applying, not someone who is simply looking for a place to land. “The motivation has to make sense,” he says. He does not expect total mastery of evolutionary theory, but he does require a sense of purpose. “They have to be clear about what they want to gain by working with me,” he says. “There has to be some sort of trajectory.”

The interview is obviously a crucial part of the hiring process (see ‘Tips for success’), but not all PIs feel like they are ready for the task. “Among new group leaders, the interview is always a conversation topic,” Chan says. “People want to know: what are the magic questions to ask?” Chan says that he simply sticks with the basics. He asks candidates about their thesis, and he asks them to clarify how much of it was done on their initiative and how much was given to them. “I’m looking for people who can learn things very quickly,” he says.

He also looks for basic congeniality — the ability to collaborate without too much friction, to engage without too much discomfort — a quality that is hard to detect on a CV. “People

## RECRUITMENT

### Tips for success

As part of a laboratory leadership course that started in 2002, the Howard Hughes Medical Institute (HHMI) in Chevy Chase, Maryland, periodically surveys its fellows and alumni to find out the key things they wished they had known before starting their first lab. “The number-one thing that comes up is choosing the right people from the get-go,” says Maryrose Franko, a science-programme manager at the HHMI’s Janelia Research Campus in Ashburn, Virginia. “Who do you get, and how do you get them? It’s important, because that person can set the tone for your entire lab.”

Spurred by the surveys, the HHMI published a book called *Making the Right Moves: A Practical Guide to Scientific Management for Postdocs and New Faculty*. The book, available for free online ([go.nature.com/xel46p](http://go.nature.com/xel46p)), includes a chapter called ‘Staffing Your Laboratory’, which covers a wide range of topics from recruitment strategies to sample questions for telephone interviews.

The lessons apply to researchers in all disciplines of science, says Franko, one of the book’s project developers. Here is some advice from the publication.

**Attracting applicants** Some of the best staff members are found through word of mouth, so let colleagues know that you are looking for good people. Include a message on your

website that you would welcome inquiries from prospective students and postdocs. You can also place an advertisement in a journal or on the website of your scientific society.

Make sure that applicants understand your vision for the lab — how it will function, what you would expect from them, and why you are excited about the science. If you see yourself as a mentor, make that one of your selling points.

**The interview** Keep the interview structured to ensure that you are asking basically the same questions of every candidate. Try to ask a variety of questions — some direct, some open-ended — to gauge their temperament, ambition and overall approach to science. Some sample questions include:

- What are your most significant accomplishments?
- What do you want to be doing in five years?
- How do you stay current in this field?
- Describe a project in which you had to work as part of a team. How did that turn out?
- What’s the biggest challenge in your current position? How are you managing it?
- Can you name a scientist you like and respect? What do you like about that person? **C.W.**



Robert Langer has built up his lab gradually.

in a lab spend 80% of their time working with each other, not with the PI,” he says.

Before a PI arranges his or her first interview, it is important to check with the human-resources department at their institution on interview policies and regulations, says Francisco Andrade, a physiologist at the University of Kentucky College of Medicine in Lexington who participated in an online seminar offered earlier this year by the American Association for the Advancement of Science on how to build up a lab. “Every university has its own way of doing things,” he says.

It is not just a matter of getting the right forms. Universities may have strict rules about what a PI can or cannot ask a potential employee. And in some countries, some questions — about age, marital status or family plans, for example — are illegal.

### AIM FOR DIVERSITY

A study published in April suggests that such personal queries might be pointless, anyway (F. M. Felisberti and R. Sear *PLoS ONE* 9, e93890; 2014). It examined the factors that predict the productivity of UK postdocs

— and found that those with children published just as often as those without. It also found that whereas postdocs from the United Kingdom were somewhat more productive at the start of their positions, researchers from other countries quickly closed the gap. “Diversity in general is a good thing in the lab,” says study co-author Rebecca Sear, a behavioural ecologist at the London School of Hygiene and Tropical Medicine. “Some people have a tendency to hire people who are like themselves. I see that tendency in myself.” She says that she sometimes has to remind herself to take a chance on workers who might have a slightly different approach from her own (see [www.nature.com/diversity](http://www.nature.com/diversity)).

Chan is proud of the global scope of his team. His current roster includes a postdoc from Croatia, a postdoc from Australia, a PhD student from Russia, a PhD student and a research assistant from the United States, and an undergraduate research assistant and an animal caretaker from Germany. Although Chan’s lab is based in Germany, they all communicate in English; he does not particularly care whether prospective lab members can

BENJAMIN TANG



speaking German, but they do have to be reasonably fluent in English.

But assessing applicants' credentials can be dicey if their accomplishments took place at far-off institutions with different grading systems. "Our department is getting a lot of applications from different countries," says Kelleher, who works at the University of Houston in Texas. "Some are like a black box. You read it and you have no idea if they're a good candidate or not." In such cases, contact with candidates as well as referees takes on paramount importance.

### CHECK EVERYTHING

Chan says that, in his experience, many applications fail to stand up to scrutiny. "A lot of CVs claim to have every skill on Earth," he says. So more than ever, it pays to be diligent and contact supervisors as well as look at the actual publication history. "In a competitive market, a lot of people will colour beyond the lines," he says. "But if you claim to have done something that you didn't really do, that's a deal breaker."

Andrade says that every detail on an application is worth double-checking. "We see incorrect information at all levels," he says. "At best, it just shows a lack of care. At worst, it's something else." And even letters of recommendation can mislead, adds Odom, who says that he puts little stock in them and calls the referees instead. "You have to speak to a human being. Even if they were truthful in their letter of reference, they may have been guarded."

He says that grades, testing scores and endorsements from past supervisors are all important, but above all, he is looking for people with a plan, especially potential postdocs. "Grad students are there to stabilize their lives and figure out what they want," he says. "If they want to be an investment banker, cool. But if postdocs don't know what they're going to do, you shouldn't hire them. Some of them may not be planning to go into the field, but they have to have a clarity of heart."

Of course, PIs must have plans of their own. Kelleher, for her part, wants to keep her lab small — hire a couple of PhD students soon, and maybe a postdoc down the road. As a postdoc, she was in a lab with just a few other people, and enjoyed that intimacy. "That will probably be my preference as a PI," she says. Her vision for her future will really start to take shape when the next person joins the lab — whoever it is. ■

**Chris Woolston** is a freelance writer in Billings, Montana.

# TURNING POINT

## Juan David Ramírez

*Juan David Ramírez, a postdoc in molecular parasitology at the US National Institutes of Health (NIH) in Bethesda, Maryland, was named a Pew Latin American Fellow in June. After the two-year fellowship, Ramírez plans to return to his native Colombia to help fight his country's endemic parasites.*



### What sparked your interest in parasites?

I come from a country with many endemic tropical diseases. Many people in my family had malaria. One had Chagas' disease. I became really interested in infectious diseases, particularly those caused by parasites. Luckily my teachers in high school encouraged my love of microbiology, and I decided to study it as an undergraduate at the University of the Andes in Bogotá.

### What made you pursue a graduate degree?

During my bachelor's, I developed a molecular test for diagnosis of Chagas' disease. When I finished that, I did a master's examining the link between genetic diversity and clinical outcomes. Only two drugs are available to treat Chagas' disease. My adviser, collaborators and I found that most of the parasites (*Trypanosoma cruzi*) were resistant to one of the two, and developed a test to determine which drug should be used in each patient. Our results helped to create a guide for treatment of the disease in Colombia. I want to do similar work on other parasites.

### Describe your graduate experience.

My adviser was supportive and let me do anything I wanted. I was an author on 18 studies on the molecular epidemiology of parasitic diseases in journals such as *PLoS Neglected Tropical Diseases* and *Acta Tropica*. We were in a good situation — we had close contact with patients and clinical metrics of the disease. I also had the opportunity to spend a year at the London School of Hygiene and Tropical Medicine. I brought parasite samples from humans, reservoirs and insect vectors in Colombia and explored the genetic diversity and reproductive mechanism of *Trypanosoma*.

### Eighteen publications seems like a lot

It was. I won the national science award as a result. I owe a lot to my supportive adviser, but I was quite focused on publications, serving as primary author on 12 studies while also providing samples or analysing data for collaborations. As long as I had interesting results, I pushed my adviser to read and correct the manuscript. I wrote so that we could submit for publication.

### How did you secure a postdoc at the NIH?

While I was doing my PhD, the Latin American Congress of Parasitology convened in Bogotá. There, I met my current adviser, Michael Grigg. He had seen my work on *Trypanosoma* markers and liked it, and was doing similar work in *Toxoplasma*. I asked about the possibility of coming to the NIH to do a postdoc, and e-mailed him when I finished my PhD. In April last year, I started a postdoc on *Leishmania* and *Giardia*.

### Describe your postdoc.

It is awesome. In Bogotá, where I did my masters and PhD, we had restrictions on resources, equipment and technology. Here, the sky is the limit. I do not have to worry about not having access to a sequencer.

### What does the Pew award mean to you?

I am the second Colombian in history to get the award and that is important to me. Research in South America is focused largely in Brazil, Chile and Argentina. Other countries have talented researchers but do not get many opportunities. The award is also important because it provides funds if I want to return to Colombia to start my own lab after two years here.

### Will you return to Colombia?

Yes. I was productive in Colombia as a graduate student and got research funded by the European Commission. I think I can still do that. I want to help Colombian science to be better appreciated and to do good work that will help to persuade the government to invest more in science. There are many other parasites I want to explore. I want to do work that has an impact on the health of my country. ■

INTERVIEW BY VIRGINIA GEWIN

# WELCOME TO THE WORLD, TRILBY FREEDOM

*It's the age of enlightenment.*

BY MARCELINA VIZCARRA

The whole cul-de-sac had been invited. Arabella lifted the visor to reveal her daughter's chubby face, fresh from her birth-year quarantine. "Her name's Trilby Libertas. It means Hat of Freedom," Arabella said. The other parents at Jared's table moaned in approval. It was a good coup in the clothing-name trend. You couldn't take two steps anymore without passing a Vesta or a Helmet, a Buckley or a Blazer. But a Trilby ranked up there with past icons like Snobia, Anavrin and Chyme for uber unique.

The party noise triggered the audio-contamination warning on Trilby's second womb. Arabella adjusted the baby's acoustic-foam headband. "Anything higher than 85 decibels can damage hearing," she announced. The parents murmured in agreement.

The kids at the indoor park lolled inside their transparent filter-suits like unhardened vegetables. When Jared was a child, the dome's climbing walls and tunnels were attempted with bare hands, the bacteria-and-sugar mucilage an inevitable, occasionally welcome, form of traction. He'd used the gunk once to grease Arabella's ponytails. The next day at school, she'd appeared with a fresh pixie cut and a zero-tolerance policy in her hand.

"Swimming lessons!" Merit's voice flared with mock-indignation. "Right! One life-guard per ten kids."

Enlightened parenting. Jared's least-favourite default conversation. Better to pitch in, though, than to allow his silence to be mistranslated throughout cyberspace as aloofness again. And end up on the apathy-police's radar. Talk about creepy fora. "We can't blame our folks," Jared said. "They didn't know any better."

"Society wasn't ready."

"Exactly. My mother was ridiculed for putting me on a leash."

"Speaking of inhibitors, what about those car seats!" Seattle said. "Remember those? Like our backs were somehow more important than our organs."

"Those harnesses caused my lactose intolerance."

A chorus of beeps alerted the parents to the passage of time, and they turned in unison to monitor their children. At least, here, with the kids safely herded, they didn't

have to suffer the child-hostile public spaces, the nanny drones following them through supermarkets and shopping centres.

"My gripe: the backyard playsets," Arabella said. "My mom used to send me outside by myself *every afternoon*."

"Me too," Seattle said. "Once, I played with cat faeces for an hour before she came outside and told me what it was. Toxoplasmosis, anyone?"

Bangle rolled over for a portion of gluten-free, nonallergenic, free-radical-blasting vitamin pulp. Jared leaned as he shifted in his chair, setting off the girl's proximity alarm.

"Christ, I had no idea I was that close." He apologized profusely to Bangle's mother who, being gracious, recorded the alarm.

"Don't worry. I know you're not a perv." She laughed. She'd checked him against the registry two weeks earlier when he complimented Bangle's dress. "Meet Trilby, Bangle. Not so close, honey. Stay behind the sensor."

Too late. The womb pulsers discharged, delivering a preliminary, non-lethal shock.

"I got bit once from my great-grandson's carrier," an old man said. One of the relics from the cul-de-sac. "Remember when you'd shock your fingers against the doorknob after walking across the carpet? It's about like that." He made a buzzer sound and poked Jared in the shoulder.

The others regarded the man with pity. He coughed into his fist. On cue, they pulled out their pocket sanitizers and masked themselves, effectively ending the chat.

Arabella narrated into her diary for the benefit of her 8,000 followers. "Having a blast. Trilby's met everyone except Panto." She raised her eyebrows at Jared. He shuffled through the apps on his diary to summon his son from the pretzel slides. The action posted online, where half a dozen people including a retired couple in Montana and Jared's mother in Florida reposted the event. Panto wheeled over.

Jared had chosen the name in panic after his rival neighbours announced their twins, Sari and Knickerbocker. At least he'd never force his son into friendships with the snobby twins. Gone were the days of the shoving, pinching,

sticky touch of childhood, the adhesion of germs and bad decisions and parental disinterest that used to form bonds between kids. Perhaps the parenting enlightenment had accomplished something after all, Jared thought.

"A toast," the elderly man said, lofting a tumbler of carbon- and politic-neutral banana splash, "to the little lady on her first birthday. Welcome to the world, Trilby Freedom."

Arabella winced.

"Hear, hear." They quaffed their juice while Trilby chewed on her crib mentor, a plush kangaroo translating their every word into Mandarin and Japanese.

"And," Jared said, rising to his feet in a rare fit of self-promotion, "to the sacrifices we've made for our children. Thanks to modern understanding, they can pursue real friendships, based on respect instead of proximity."

His neighbours blinked and paused, then, pleased with Jared's epiphany, self-attributed it in their diaries. Panto hugged himself, signalling Bangle to hug herself in response. They finished with an air high-five, delivered, by the push of a button, with a cymbal clash.

On cue, the sandwiches arrived in their own dome. The server detailed the chicken salad's previous incarnation as animal and vegetable and offered to list toxin exposure for each ingredient. Arabella waved away the server. "We're not fascists here. I mean, we survived our parents' menu."

"Back in the day, before free-range food."

"Honey, your dinner was clinically depressed. Enjoy!"

Jared relaxed. Free-range food, air and sunshine. At last, something everyone could agree upon. ■

Marcelina Vizcarra lives in the Midwest with her family.



JACEY

➔ NATURE.COM  
Follow Futures:  
@NatureFutures  
go.nature.com/mtoodm